RICE UNIVERSITY

The female-to-male transsexual voice: Physiology vs. performance in production

by

Viktória Papp

A Thesis Submitted
in Partial Fulfillment of the
Requirements for the Degree

Doctor of Philosophy

Approved, Thesis Committee:

Katherine Crosswhite, Chair
Assistant Professor of Linguistics

Nancy Niedzielski
Associate Professor and Chair of Linguistics

Paula Sanders
Professor of History and Dean of Graduate and Postdoctoral Studies

Monica McHenry
Associate Professor of Communication Sciences and Disorders, University of Houston

Houston, Texas
June, 2011

ABSTRACT

The female-to-male transsexual voice: Physiology vs. performance in production

by

Viktória Papp

Results of the three studies on the speech production of female-to-male transgender individuals (transmen) present phonetic evidence that speech produces the transmen by what I termed triple decoupling.

Transmen successfully decouple gender from biological sex. The results of the longitudinal studies exemplified that speakers born and raised female do not necessarily need to have a female voicing source or filter function. Both qualitative changes can be achieved (to different degree) by bringing exogenous testosterone into the system that virilizes both source and filter over time. Moreover, the cross-sectional study showed that articulatory gestures can be modified to move the acoustic targets towards a gendered target one is striving to present.

The acoustic manifestations of transmen with different partner attraction offers the next type of decoupling, that between sexual orientation and gender identity. The results of the cross-sectional study imply that female-born individuals attracted to men do not necessarily have to identify as women. They can opt out of this self-identification by selectively adopting features associated with the gay cismale speaking style. This is suggested by the fact that sexual orientation was found to have a significant effect on the durational and spectral quality of fricatives /s/ and /s/, formant values and sentential pitch range.

Finally, the longitudinal studies provide evidence for the third type of decoupling, which comes in the form of gender breaking free from physiology. The recurring "reverse J-pattern" of both the transitioning source and filter, as well as the mean fundamental frequency raising above the pitch floor illustrate the fact that transmen do not feel obliged to sound as masculine (as low-pitched and "low-formanted") as testosterone enables them to. This final type of decoupling also serves to demonstrate that many transmen decidedly do not opt in to the binary system of sex / gender – even though they are physiologically able to do so.

Although LGB speaking styles have been investigated before, this dissertation is the first to discuss a number of acoustic descriptors specifically in transmen's speech and place them into the context of hormone treatment, sexual orientation and disclosure status.

Acknowledgements

My most heartfelt gratitude is to the men associated with the Transgender Foundation of America (TFA) for inviting me to witness a beautiful part of their lives over the past four years. My relationship with these people is based on love and respect more than anything else. Both despite and because of our substantial cultural and experiential differences, my long-term relationships with my transmen friends have changed my character and my sensibilities forever. Any merit that this dissertation has is dedicated to them, because they truly made it possible. Therefore, even though this study was not written directly for the transmen at the TFA as an audience, it is entirely for them as my friends and collaborators, and I offer it as a sign of my respect for their courage, determination, and sense of humour.

I would like to make special mention of two individuals who have deeply enriched my life as well as my research. Stacey Colt Meier was the first person who transitioned in my life and he shared not only his transitioning but later his research as well with me. I count myself as very fortunate to have been in the right place, at the right time, to have had the opportunity to start this project with him and build the friendship I have with him. He and Lou Weaver have shown great enthusiasm for and patience in not only documenting their speech but introducing me to the gender-variant side of Houston, finding further subjects, hand-delivering them to the labs if necessary, and making me a member of their extended family.

Among the men I met at the TFA, Trey Johnson, Mitch Sellers, Jess Hammons, Chase Angel Fortuno and Eric Brown have been especially good-humoured and patient friends in and out of the lab and I consider each of them a role model for me in how to gracefully travel through and around genders. To every transman, transwoman and gender-queer individual in Houston, who has shared their friendship with me, my gratitude is equally sincere.

The longitudinal data gathering part of this project would not have been possible without the support of a host of other graduate and undergraduate students in the Linguistics Department. I would like to say thank you for covering some of my sessions over the four years to Linda Lanz, Chris Koops, Anne Marie Olivo, Cassandra Pace, and Natalie Weber. Without you, no trend line would be as smooth as it is. Additional thanks go to Linda, Cassandra, and Natalie again for helping me make not only trend lines, but also my prose smoother.

The cross-sectional study would not have been possible without the generous help of Monica McHenry who gave me free reign of her lab at the University of Houston and much support for the duration of the entire project.

I am also grateful to my family for their support and encouragement over the years, most especially in those moments when they were (admittedly rightly) sceptical about the advisability of my plans.

I am very grateful to the members of my dissertation committee for their contributions to this work. Each of them has given me much-needed support, encouragement, and feedback, and each has provoked me to think more carefully and clearly about various issues of mutual interest. Katherine Crosswhite and Monica McHenry have been the greatest mentors in laboratory work a graduate student can ever wish for and both of them have been a source of inspiration regarding the interweaving my research findings with phonetic and clinical theory. Nancy Niedzielski has in her own way also pushed me toward a deeper engagement with the foundational perspectives on gender and language. I feel honoured to have worked with each one of these scholars.

I am deeply grateful for the financial support I have received over the years that made the data gathering process and the dissertation writing possible. Besides the donors behind the graduate stipend from Rice University, I owe much gratitude to

Richard and Jackie Schmeal who for five years allowed me to stay at their garage apartment on Mandell Street. Their generosity gave me not only a home but also the financial freedom to focus my energies on my studies, research, and intellectual development in general.

Over the past year, I have had the privilege of conversing with an exceptional fellow graduate student, Bethany Townsend. I would like to thank her for staunchly standing her ground in our endless quantitative-qualitative battles and making me reconsider my research from a qualitative point of view. The final chapter in this dissertation bears the signs of those late night conversations and is offered to her in hope of future collaboration on this community. May there be vegan gnocchi too.

I do not take for granted all of the support from other people – be it logistic, intellectual, or emotional in nature – that has made this dissertation possible, and my life in Houston so rich. Therefore, I end here by offering my thanks to everyone not already mentioned here who has contributed to my work and life. I hope you can feel my gratitude.

Contents

Illustrations

Tables

Chapter 1

Introduction

1.1 Overview of the dissertation

This dissertation explores the performance of speaker gender and sexual orientation in speech production. Previous research on gendered talk has shown that female and male speakers speak in different fundamental frequency ranges. Their formant values and the acoustic correlates of their consonants differ as well. Theories of vocal tract normalization, also referred to as vowel or speaker normalization, so far have strived to factor out physiological differences in order to make sociophonetic comparisons. The current dissertation takes multiple approaches to analyse how physiology and performance of speaker gender and sexual orientation shape the acoustic signal in the speech of female-to-male transgender and transsexual individuals.

This dissertation fills a gap in our existing knowledge about the vocal changes that female-to-male transsexuals (FTMs) go through as they transition from a female-bodied individual to a (more) male-bodied individual with the help of hormone therapy. Although a rich literature exists for the voice and speech issues of male-to-female transsexuals (MTFs), relatively few works explore either the transitioning or the established FTM voice. This particular imbalance in the literature reflects the wide-spread clinical assumptions of speech-language pathologists and endocrinologists that testosterone therapy masculinizes the voice of FTMs to the extent testosterone masculinizes the voice of adolescent biological males.

However, Gorton, Buth, and Spade (2005) and van Borsel and his colleagues (van Borsel, De Cuypere, Rubens, & Destaerke, 2000; Adler & van Borsel, 2006) warn FTMs about potential problems with using hormone therapy. Gorton et al. recommend that "professional or amateur singers and speakers should be warned that frequently voice changes occur that may be significantly detrimental to vocal performance" (Gorton et al., 2005, p. 59). This is challenged by the musicologist Constansis (2008) who coached a number of vocally trained transmen after their transition, with varying results. Some of his subjects, including himself, achieved an impressive fundamental frequency range that was larger than before coaching, although he does not provide evidence of the original size of the ranges before testosterone treatment. The voices of some transmen under his coaching admittedly changed little or not at all. Constansis, a musicologist, refers to this vocal phenomenon as "entrapped FTM voice," and characterizes it as "lacking contour and colour" and displaying permanent hoarseness and low intensity. The two individuals in Constansis' study who had "entrapped" voices had incompletely descended larynges (between the characteristic female and male positions) that were only slightly larger than those of sopranos. He also noted that the voice descriptors of these two transmen did not change even during / after voice therapy and they exhibited constant hoarseness and inability to access and control certain vocal areas.

The goal of the first longitudinal study (described in Chapter 3) in the dissertation is to follow five FTM individuals over the course of their first year on testosterone and observe the ensuing changes in mean fundamental frequency as well as fundamental frequency range in order to establish the magnitude and time course of the changes in individuals.

It is established in the literature that the fundamental frequency of transitioning FTMs decreases and the vocal changes that manifest early on in the treatment do not reverse. However, only a handful of studies have documented this process (Oates & Dacakis, 1997; Söderpalm, Larsson, & Almquist, 2004; van Borsel et al., 2000; van Borsel, De Cuypere, & van den Berghe, 2001). The surveyed literature generally agreed that testosterone therapy was sufficient in creating a pitch that allows "passing" as a male (Andrews & Schmidt, 1997; Davies & Goldberg, 2007; de Bruin, Coerts, & Greven, 2000; King, Lindstedt, Jensen, & Law, 1999; Oates & Dacakis, 1997; Söderpalm et al., 2004). However, the studies offering this conclusion do not provide empirical verification of this widely held idea. The only vocal satisfaction survey (van Borsel et al., 2000) available demonstrated that a four of the sixteen FTMs who had taken testosterone for at least 12 months were sometimes perceived as female over the phone, and five of them expressed interest in speech therapy to further the male perception of their voice.

The claim of the non-reversibility of changes brought on by testosterone is tested in both longitudinal studies.

Interpreting the existing studies is complicated by the fact that the longitudinal studies on FTM speech production (Damrose 2009; van Borsel et al. 2000) use very few subjects, and also employ incorrect methods to estimate vocal ranges, changes in pitch, and voice quality descriptors, such as jitter, shimmer, and noise-to-harmonic ratio. The researchers in both studies subtracted Hz values without a reference frequency to calculate vocal ranges and changes in pitch. This method has some problems. A preferable technique is to use a perceptual scale to express changes, such as the semitone (ST) scale. Use of the incorrect estimation procedure resulted in post-testosterone values for jitter, shimmer and noise-to-harmonic ratio that were near-pathological (Damrose, 2009). This approach does not take into consideration that these descriptors naturally increase in a deepening voice when expressed in percentages.

This dissertation also expands on previous research by collecting and examining data at a larger resolution than in previous studies (Damrose, 2009; van Borsel et al., 2000). Instead of six or twelve measurement points over a year, my developmental study documents the changes of transitioning voices with about 30 measurements over a year, which makes the data more comparable with available normative data on the vocal changes of biological males during adolescence (Pedersen, 2008).

Another lacuna in previous research on FTM voices concerns changes in formant values. To my knowledge, there is no documentation of changes in formant values or the length of the vocal tract, with the possible underlying assumption being that they do not change if the speaker starts testosterone therapy past puberty. As Beck (1999) indicates, the growth of the mandible is highly sensitive to a variety of factors. Bevis, Hayles, Isaacson, and Sather (1977) document that the mandible also responds more to growth hormone than most other bones. I hypothesize that the mandible may also be more responsive to testosterone. Since appreciable changes are anecdotally noticeable in the angle of the mandible and the angle of the jaw line after one year on testosterone, the second longitudinal study (described in Chapter 4) examines changes in the formant values of five FTMs.

Following the analysis of longitudinal changes in the speech production of FTMs, in Chapter 5 on page 103 the speech production of a larger group of transmen is carried out in a cross-sectional study. The lack of research on the production and perception of FTM voices without visual cues leads the FTM community to rely on self-introspection and anecdotal advice in speech production, which may lead to voice pathologies. The anecdotal observation that many FTMs sound like gay males or androgynous individuals is examined with the help of an extensive acoustic analysis of transmen's voices that includes pitch, formant and fricative measures.

In the final chapter, the results of the acoustic analyses are integrated with a Butlerian performative approach to gender and sex, which is in turn used to shape the debate about the connection between identity and linguistic variables, as well as the relationship between physiologically pre-disposed vs.performed features in speech.

1.2 Definitions and terminology

Before discussing who exactly transmen are, it is important to distinguish a range of terms either used as labels or categorization tools for those labels, such as gender, sex, identity, orientation, performance, transgender, transsexual, etc. As nearly all of the available definitions cross-reference more than one of these terms, a certain degree of circularity is unavoidable. The terminology laid out below is a set of working definitions and is not intended to capture the ongoing extensive social, political, medical, legal and psychological debate about each and every one of the terms.

I use the term biological sex to mean the manifestation of the combinations of genetic (chromosomal) and hormonal information packaged in physical bodies. The two main types of human sex are male and female, with other known types exhibiting partial or mixed characteristics, such as individuals on the intersex continuum.

Gender is a social construct that, depending on the scientific approach, is used to distinguish between male / masculine vs.female / feminine traits based on biological sex, gender role, gender identity and presentation, as well as sexual orientation (Eckert & McGonnell-Ginet, 2003). Gender is usually based on sex, but it exaggerates sex, and extends it to domains where it is not relevant (Eckert & McGonnell-Ginet, 2003; Jordan-Young, 2010).

Having originally meant grammatical gender only, a gender-sex distinction grew out of social-psychological work on gender-non-compliant and gender-variant individuals starting in the middle of the 20th century. Along with that came the notion that Judith Butler (1999) terms "performative" gender: "Gender is the repeated stylization of the body, a set of repeated acts within a rigid regulatory frame which congeal over time to produce the appearance of substance, of a 'natural' kind of being" (1999, p. 33). In this light, gender ceases to be something that an individual has or is, and becomes something that one does. From that follows that we should understand "woman" and "man" as open-ended terms and terms "in process, a becoming, a constructing that cannot rightfully be said to originate or end [... and] it is open to intervention and resignification" (Butler, 1999, p. 43).

Kulick (2003) gives succinct definitions of performance, performativity and performative approach. Performance is something a subject does and it is one dimension of performativity. The linguistic performance framework, such as the classical variationist approaches, has the tendency to see language in relation to identity, which is sometimes presented as more or less conscious claim-staking of a particular sociological position. This framework asks and answers the question "Who says it?". This approach tends to start or end with a pre-discoursive identity.

Performativity is the process through which the subject emerges. This framework concentrates on identification, a process which is concerned with the operations through which the subject is constituted. This framework asks and answers questions more along the line "What does saying / not saying it produce?". Unlike the performance framework, the performative approach does not start or end with identity, but it examines the processes through which some kinds of identifications are authorized, legitimate and unmarked and some are unauthorized, illegitimate and marked. It also represents the shift in research focus from the sociolinguistics of identity to the sociolinguistics of identification.

Twenty years after Butler's work *Gender Trouble*, while the academic community largely treats biological sex and gender as separate constructs, contemporary mainstream US society, including clinical practice, typically views gender as a function of sex. From the academic vantage point, Judith Butler's work (1999) introduced further complexities into the performative model of gender:

> "If the immutable character of sex is contested, perhaps this construct called 'sex' is as culturally constructed as gender; indeed, perhaps it was always already gender, with the consequence that the distinction between sex and gender turns out to be no distinction at all" (Butler, 1999, p. 10–11).

Mikkola (2008) reconceptualizes Butler's views the following way: Since both sex and gender are the same, and gender is socially constructed and performed, sex is performed as well.

Where does the idea that sex is performative come from? Butler's (1999) view is that "sexed bodies" are also performative in as much as they exhibit "no ontological status apart from the various acts which constitute [their] reality" (p. 173). In Mikkola's interpretation this implies that female and male bodies do not have independent existence and that if gendering activities ceased, so would sexed physical bodies. Butler's account of gender performativity also explicitly renounces essentialist views of identity. She refuses to account for genders as bearers of various essential and accidental attributes. In her view (Butler, 1999, p. 24), feminists who consider genders socially constructed in an essentialist way define "women" as females with feminine behavioural traits, being heterosexuals whose desire is directed at men, and "men" as males with masculine behavioural traits, being heterosexuals whose desire is directed at women. This feminist practice creates "intelligible genders" if individuals exhibit all this chain of traits in a coherently linked manner, i.e.,where sexual desire follows from sexual orientation that in turn follows from feminine / masculine behaviours thought to follow from biological sex. This, however, creates only two socially sanctioned genders, and multiple socially unsanctioned ones. In Chapter 6 on page 153 I discuss transmen's performance of sexual orientation / partner attraction and gender in the light of a hybrid performance-performative model.

An "intervening variable" in the above definitions is sexual orientation. According to Diamond (2003, p. 174), a person's sexual orientation refers to their "relatively stable tendency to seek sexual partners of the same gender, other gender, or both genders," while sexual orientation identity is the label ascribed to these partner attractions. Though these concepts are a useful tool for the majority of the population, they are based on self-identification by the individuals and assume unchanging sexual and gender identities on the part of all individuals in a particular relationship.

An individual can be defined as transgender if their "gender-related identification or an external presentation either violates conventional conceptualizations of "male" or "female" or mixes different aspects of male and female role and identity" (Diamond & Butterworth, 2008, p. 365). Since many transmen originally identify, for lack of better available terminology, as lesbian women, after transitioning into a more male-bodied individual— assuming no change in partner preference— the same individual would be perceived as a straight male. As a result, if his partner's gender identity stays stationary, the partner will be perceived as a straight female. The two of them will constitute a "straight couple," a term they may or may not identify with. As we can see, the sheer existence of transitioning and the reliance on self-identified labels might cause terms such as "sexual orientation" to inadequately capture the changing life experiences of the subject population of this dissertation. In order to operationalize this important characteristic, instead of canonical sexual orientation labels, this dissertation uses the concept of partner attraction. While this practice is admittedly equally subject to complications, it reduces the level of abstraction in characterizations.

Transsexuality can be most broadly defined as not identifying with the gender assigned to one at birth, or more specifically as taking hormonal, surgical, or legal steps to live as the sex opposite from the one assigned at birth. Despite this simple sounding definition, precisely who identifies as transsexual and / or transgender and who is socially, medically or legally recognized and accepted as such is a highly politicized matter. Recent medical-legal discourse typically draws a line between transgender and transsexual individuals by requiring hormonal or surgical modification to qualify as transsexual. Medical discourse also pathologizes cross-gender identification. Although "homosexuality" has not been listed as a mental illness in the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV-tr, American Psychiatric Association (2000)) since 1973, the current manual still includes the diagnosis of "gender identity disorder" (GID), defined as a "strong and persistent cross-gender identification" (2000, p. 581).

Acronyms such as LGBTQI (Lesbian Gay Bisexual Trans* Queer / Questioning Intersex) and various other reincarnations, commonly referred to as the "alphabet soup," imply that transgenderism and transsexuaslism are types of sexual orientation when they are in fact gender identities. Nothing proves this more than the observable heterogeneity in partner attraction within the transsexual and transgender population. Although the current version of the DSM-IV states that transmen who are sexually attracted to men are exceptions to the general rule (American Psychiatric Association, 2000), Meier (in-press) found that over half of his 600+ sample were attracted either to males only or both males and females. Additionally, he found that this population is also highly heterogeneous in gender identity, gender expression, gender and sexual fluidity, surgical body modifications, legal and marital status and hormone use.

More specifically, transmen or female-to-male individuals (FTMs) are individuals who were assigned a female birth sex, but currently identify as male. As already delineated in the transgender-transsexual distinctions, FTMs may or may not employ medical, surgical, hormonal or legal acts to be recognized as male members of the society. Some of those whom society consistently perceives as the intended or preferred sex / gender may consider themselves men and no longer identify as transsexual (see Meier (in-press)). Many may not disclose trans history voluntarily and consider their transition as corrective surgery or treatment of a congenital birth defect.

The language in the area of transgender / transsexual experience is constantly evolving, and the self- and other-applied taxonomy coming out of the trans and the academic community is not only changing, but displays geographical and racial variation.

For example, various populations in many African American communities use "stud" to self-label either lesbian identified females, gender variant individuals or transmen. Similarly, blogs, mailing lists, in online academic communities or online surveys and the trans community itself experience continual flame wars about the choice between "trans man," "trans-man," and "transman." My choice to write it as one word was not motivated by any salient ideological / semantic choice other than that it was the prevailing spelling at the beginning of my longitudinal data collection in 2007.

In this work, following community politeness norms, I use *he, him, his* and *man / men* to refer to transmen once they declared male or transgender identity to me, regardless of their gender presentation or current gender role in society. I also frequently make use of the "singular they" when the gender of the individual in question is not pertinent.

I also use both "transgender" and "transsexual" as terms, but not interchangeably. Without wanting to rely on the medical (or medicalized) dividing line between the two communities, which largely overlap spacially, temporally and experientially, the distinction serves a useful purpose by highlighting cross-gender identification in both groups, but testosterone use only in the latter. This does not imply that exogenous testosterone was absent from the transgender participants in my studies, but shifts the focus to the social nature of living as a man. When I refer to the entire community that encompasses both transgender and transsexual individuals, I use the umbrella term "trans."

Finally, in order to avoid the so-called "othering," or defining individuals whose birth sex matches their gender identity with a negated term (i.e.,"non-trans"), following contemporary practice established in the 1990s, I refer to them as "cisgender" individuals*.

*According to the Oxford English Dictionary (J. Simpson & Weiner, 1989), Originally a Latin preposition, *cis* meant "on this side of" and was used as the antonym of the prepositions *trans* "across' and *ultra* "above." Later, more prefixal, usage included spatial and, less frequently, temporal uses as well: cisalpine "on this side of the Alps," cisatlantic "non-trans-Atlantic (e.g., trade)," Cisjordan "West Bank," cis-Elizabethan "pertaining to the era since Elizabeth I."

Chapter 2

General literature review

2.1 Introduction

In this chapter, I present an overview of gendered vocal tract anatomy and sound production in the source-filter theory of speech production framework. I review the developmental changes that take place in puberty, summarize the available literature on ageing and the ossification of the larynx, and describe the gendered normative physiological structures for adults. A section on the previous research done on FTM voices concludes the literature overview.

2.2 Source-filter theory of speech production

In the experiments comprising this dissertation, the fundamental frequency, formant and fricative characteristics of speech are assumed to be some of the main correlates of speaker gender and sexuality. To capture the acoustic differences in the speech signal, the phonetic analysis in this work is based on the acoustic source-filter theory of speech production.

The acoustic differences that allow listeners to determine phoneme identity amongst vowels and consonants are often explained by the source-filter theory (Fant, 1975). The "source" is the sound spectrum created by airflow through the glottis, which varies as the vocal folds vibrate. The "filter" is the vocal tract itself, the shape and the length of which are controlled by the speaker. As the source waveform passes through the vocal tract, the filter characteristics modify the source by passing certain frequencies and attenuating (or dampening) others. In each production study the source and the filter characteristics of transmen will be investigated separately.

2.3 The gendered vocal tract anatomy

2.3.1 Laryngeal structures and the control of phonation

Since the acoustic characteristics of the source are to a great extent based on the physiological characteristics of the human larynx, this section opens with an overview of the laryngeal anatomy and mechanics.

The framework of the larynx consists of four different cartilages: the epiglottis, thyroid, cricoid, and arytenoid. The thyroid and cricoid cartilages are connected by the cricothyroid joint, while the arytenoid and cricoid cartilages are connected by the cricoarytenoid joint. The movement of the cricothyroid joint changes the length of the vocal folds. Movements of the arytenoid cartilage on the surface of the cricoarytenoid joint contribute to the abduction-adduction of the vocal folds. The main movement of the cricoarytenoid joint is a rotation (abduction-adduction) of the arytenoid cartilage around the longitudinal axis of the joint. Other possible movements of the arytenoid are a small degree of sliding motion along the longitudinal axis of the joint and a rocking motion around a fixed point at the attachment of the posterior cricoarytenoid ligament (Leden & Moore, 1961). For further details of the anatomy of laryngeal structures, see Gauffin and Hammarberg (1992) and Titze (1994).

Movements of the cricothyroid and cricoarytenoid joints are controlled by the intrinsic laryngeal muscles. Elongation and stretching of the vocal folds is achieved by contraction of the cricothyroid muscle. Movements of the arytenoid cartilage and the resultant abduction-adduction of the vocal folds are controlled by the abductor and adductor muscles. The posterior cricoarytenoid muscle is the only abductor muscle, while another three (the interarytenoid, lateral cricoarytenoid, and the thyroarytenoid muscle) are the adductor muscles. Contraction of the cricothyroid muscle may also result in a small degree of glottal abduction. The vocalis muscle, which is the medial part of the thyroarytenoid muscle, contributes to the control of the effective mass and stiffness of the vocal folds rather than to abduction-adduction movements.

While the larynx develops in line with the rest of the body throughout childhood, between the ages 10-15 there is a sudden increase in the rate of laryngeal growth and size. Girls' voices mature due to an enlargement of the larynx that is consistent with general body growth. Boys' voices, however, drop about an octave, and the vocal folds double in length (Greene & Mathieson, 1989). At the same time, the angle of the thyroid cartilage decreases in males which produces a characteristic notch, known as the Adam's apple. Greene and Mathieson (1989) noted that contralto girls' vocal folds in contraltos grow from a prepubescent 15 mm to a postpubescent 17 mm, which is about the minimum length of the grown male tenor vocal folds as well. The voice quality differences between a female contralto and a male tenor, both with about 17 mm long vocal folds, is due to differences in the size and shape of the larynx, pharynx and chest. As a comparison, Greene and Mathieson found that in the bass voice the vocal folds grow to about 23 mm.

There seems to be little disagreement about average adult vocal fold length, which is usually reported to be anywhere between 17.5 and 25 mm in males, and between 12.5 and 17 mm in females (Titze, 1994). Hirano, Yoshida, Yoshida, and Tateishi (1985) report slightly smaller adult measurements, suggesting a total vocal fold length of 17-21 mm in males, and 11-15 mm in females. The effective vocal fold length, the length of the part of the vocal folds that is actually in vibration, is estimated to be 16 mm for males and 10 for females (Titze, 1994). Hirano's results, based on a Japanese population, may well reflect geographical differences in laryngeal dimensions which derive from genetic variation between populations.

The relative proportions of the membranous and cartilaginous parts of the vocal folds are usually reported to be broadly similar in both sexes, with the membranous part constituting about two thirds of the total vocal fold length in adults. Hirano et al. (1985) have, however, shown that the ratio of cartilaginous to membranous portions of the vocal fold changes throughout childhood, and that there is a slight gender difference in the adult ratio. In newborns, the cartilaginous portion of the vocal fold constitutes slightly less than half the total length of the vocal fold, but disproportionate growth of the membranous portion results in a relative and absolute increase in length of the adult membranous vocal fold. This is slightly more marked in boys, so that in adult males the membranous portion of the vocal fold constitutes a little more than two thirds of the vocal fold length.

These differences suggest that the vocal source is physiologically naturally highly sexually dimorphic. As such, transmen are predicted to have to bridge this dimorphy either physiologically or behaviourally. Their success in doing so is documented both developmentally and cross-sectionally in the production studies, Chapter 3, Chapter 4 and Chapter 5, respectively.

2.3.2 The gendered vocal tract

In the previous section we have seen a characterization of the anatomy and the mechanics of the voicing source. The source waveform, however, is not directly emitted into the environment but is first filtered by the vocal tract. In this section the developmental changes are summarized that lead to emergence of a gendered vocal tract.

In her anatomical review, Beck (1999) summarized that the growth of the mandible (lower jaw) seems to respond to growth hormones more sensitively than most other bones and she suggested that it may also be more responsive to testosterone. She documents that the angle between the ramus and the body of the mandible is gradually reduced from about 140 degrees in infancy to 120 degrees in adulthood. Beck also (1999) asserted that the greatest contribution to overall facial growth at the time of puberty is made by the mandible. During this period, most growth continues in the ramus, but there are also marked increases in the length of the body and the vertical distance between the chin and the incisors.

During the time leading to adolescence, pharyngeal volume increases dramatically as the larynx adopts a lower position in the neck. At puberty there is another period of pharyngeal enlargement, which is more marked in males, as the larynx descends further. To illustrate the differences, Stemple et al. (2000) document that the larynx comes to rest in front of vertebra C6 in adult women but vertebra C7, i.e. an entire vertebra lower, in adult men. As this descent is accompanied by general facial growth, it creates an expanded vocal tract, which contributes further to the change in resonance characteristics.

In sum, the most significant sex-related differences evident by early adulthood concern the overall size of the vocal apparatus, the relative size of the larynx and the pharynx, and the relative proportions of the resonating cavities. Both sexes show some growth in vocal tract size during this period, and full maturation of the larynx will influence the range of phonation available to each individual. Growth of the vocal apparatus at puberty in girls can be seen mostly as growth of the pre-pubertal vocal apparatus in line with general body growth, but in males there are significant changes in the relative proportions of the vocal apparatus. The size of the male larynx increases rapidly and disproportionately, and the pharyngeal cavity increases its size relative to the oral cavity.

As we can see in the vocal maturation of biological males, the two changes – the growing vocal folds and the lengthening vocal tract (by virtue of the larynx descending) – happen together causing the characteristic pitch drop and voice quality change. In contrast with this, FTMs are hypothesized to go through a change in which the vocal folds are growing but are housed in an established and therefore less pliable laryngeal structure (Constansis, 2008). While Constansis (2008) does not explicitly state that transmen's sources and / or vocal tract resonances do not achieve cismale values during the hormonal transitioning, his work entails such a conclusion. This underlying assumption in trans-related research will be examined in detail in Chapter 4.

2.3.3 Ossification of the larynx

During ageing, laryngeal cartilages are naturally subject to calcification, with important consequent changes in elasticity of the cartilages. The age of onset of calcification varies considerably. It may begin in men in their thirties, but the thyroid cartilage may still be unaffected in some 70-year olds. In women, ossification generally begins later and is less extensive (Greene & Mathieson, 1989).

The larynx is composed of cartilages the ossification of which is highly variable across individuals. Quain, Sharpey-Schäfer, Symington, and Bryce (1914) notes in an anatomical review that ossification starts in the 20s in the thyroid and cricoid cartilages and a few years later in the arytenoids. Quain et al. state ossification starts earlier and progresses faster in males than females. Jurik (1984) further notes that in his sample of 64 patients over the age of 34, the ossification of the thyroid and cricoid cartilage occurred less frequently in females than males, while the arytenoids were ossified or calcified in all females and in 94% of males. In a radiograph measurement study of 359 patients, Mupparapu and Vuppalapati (2005) found that the thyroid cartilage was ossified more frequently than the cricoid cartilage, and that these ossifications were virtually non-existent under the age of 20. Their study reinforced the finding that males show more laryngeal ossification than women and that ossification progresses with age. Garvin's (2008) forensic analysis of laryngeal ossification in 105 isolated laryngeal structures (ages of death between 15 and 89) revealed that although a consistent sequence of ossification exists, variation in timing does not allow accurate age range estimates. Garvin (2008) also notes that regardless of ancestry (black or white) females never exhibit complete ossification in the laminae, cranial branches, or anterior midline tongues. For a diagram of the individual regions of the larynx see Figure 1 in Garvin (2008, p. 1024); for the observed sequence of ossification see Figure 2 (Garvin, 2008, p. 1025). The ossification patterns show extreme variability, with some individuals up to the age of 80 years lacking any ossification in the cricoid cartilage and commencement of ossification as young as 25 years, and minimum age of complete ossification, again, as young as 26 years of age.

In sum, at some point, the cartilages of the larynx begin to calcify and lose their elasticity. This process typically starts after the second decade, and there is evidence that the cartilages may still be elastic in the seventh decade and be already ossified in the third decade. Female cartilages ossify slower and the changes progress less far than in men.

The impact of ossification is seen in two different lights in the literature. Greene and Mathieson (1989) assume that lack of ossification is part of the reason why some singers and dramatic actors retain their voices into old age, implying that ossification is detrimental to the voice. Constansis (2008) on the other hand attributes better projection to the support of the ossified cartilages in men, implying that ossification enhances stability.

As the FTMs receiving testosterone therapy in the present experiments are between the ages of 20 and 55, the larger degree of ossification may play a role in the vocal stability and voice quality of FTMs transitioning in their 40s and later which may negatively influence the amount the larynx can grow. These issues will be investigated in detail in Chapter 5.

2.3.4 The hormonal regimen of FTMs

Regimens to change the secondary sex characteristics of FTMs follow the general principle of hormone replacement treatment of male hypogonadism. In order to masculinize (or virilize) the biologically female body, testosterone or non-testosterone hormones are administered. For an overview of non-testosterone hormonal therapies, such as the use of e.g., Depo-Provera, the use of andro pro-hormones and GnRH agonists, see Gorton et al. (2005, p. 23–25). Testosterone can be injected or administered transdermally, subcutaneously, orally or sublingually. As the half-life of testosterone is about 70 minutes, constant replenishing is necessary for life. The ideal testosterone intake would be a constant 4-9 mg / day dose. This is approximated by weekly or fortnightly intakes of depot drug formulations that slowly release and therefore extend the action of the drug. In the US the two most commonly used injectable testosterone esters are testosterone cypionate (Depo-Testosterone) and testosterone enanthate (Delatestryl) which are therapeutically interchangeable for practical purposes (Gorton et al., 2005). Other superior depot drugs mixing shorter and longer acting testosterone providing more evenly distributed hormone levels, such as Sustanon and Nebido (administered in 3-monthly intervals) are more expensive and less available in the US (Gorton et al., 2005).

After the administration of testosterone cypionate and enanthate, serum testosterone levels peak within 2 to 5 days and reach baseline levels in 10-14 days. This fluctuation of serum levels over the fortnightly window causes fluctuations in mood and energy in transmen. Some transmen find that shortening the dosage interval to about a week effectively mitigates these problems (Coty Evans, p. c. 2008).

The amount of injected testosterone depends on the preferences of both the patient and the clinician. The desired speed and intensity of transitioning, the appearance of desired and undesired effects, as well as health considerations, such as possible surgeries and illnesses, may lead to adjustment of the levels either way. Constansis (2008) reports that in his experience most FTMs start with the highest recommended intake, partially due to their desire to transition as quickly as possible and partially due to lack of individualized patient care. Once a transman achieves the desired degree of masculinization, no absolute titration (e.g., the normal biological male's range of 290-900 ng/dl) but a maintenance level is targeted. The highest achievable FTM serum level is reported to be around 500 ng/dl (Constansis, 2008).

2.4 Research on FTM voices

According to Adler and van Borsel (2006), communication or speech therapy does not seem to be a common approach with FTMs. Among medical practitioners, the opinion prevails that the pitch drop caused by testosterone is enough to allows FTMs to pass as male. This ties in with Kulick's view of the difference between the medical care transmen and transwomen have access to. According to Kulick (1999), there is an absence of literature and services offering advice to FTMs on how to talk like men. Kulick sees this as an ideological fact as much as a physiological one. This he found telling about ideas and practices of masculinity and femininity. Why aren't there more research programs, literature and services offered to FTMs? The latest available clinical manuals on transgender / transsexual communication, (e.g., Adler, Hirsch, & Mordaunt, 2006) seem to support Kulick's view. The manual describes being a man as not only self-evident but achievable by weekly testosterone shots, whereas being a woman is described as a complicated set of procedures that require careful adherence to detailed, explicit instructions about how to walk, talk, gesture, cough, sneeze, sit, eat, dress, move, and display affect. Kulick also brought a parallel from the area of body modification in the trans community. FTMs typically undergo surgical interventions to create a male contoured chest (a.k.a top surgery, or mastectomy), in rarer cases to remove the female reproductive organs, and in yet rarer cases to (re-)construct genitalia to become sufficiently male bodied. In stark contrast with this, many MTFs spend years going through procedures and operations, such as breast augmentation, lip augmentation, face-lifts, rhinoplasty, chin reduction, jaw realignment, brow shaves, cheek implants, false rib removal, chemical peeling, tracheal shaving, electrolysis hair removal, and laryngeal surgery. Being a woman, both in cultural models and in transsexual practice, he concludes, requires effort, advice, and help. Being a man, however, seems easy. But is this really true?

Adler and van Borsel (2006) reported that many FTMs reach speaking fundamental frequencies on par with normative male values. They admitted, however, that the post-transition voice is not always without problems, and the hormonally induced voice change is not as straightforward as previously thought. Van Borsel and his collaborators (Adler & van Borsel, 2006; van Borsel et al., 2000) reported a self-evaluation survey of 16 FTMs who reflected on their vocal transitioning, including the extent and the time course of the changes, as well as self- and other-perception of the changes. Out of their 16 subjects, 14 observed voice change starting at the beginning of the hormone treatment, three commented on diminished singing abilities, and subjects perceived the onset of vocal changes anywhere between a few days after the first testosterone shot to greater than one year on testosterone. Five subjects hoped for faster changes and three of them expected more pronounced changes. Two reported no change from an already low voice. Out of the 16 subjects, 13 were pleased with their voice, two wished for a deeper voice and one felt strained speaking at a low pitch. After a minimum of one year on testosterone, three transmen were still occasionally addressed as women in face-to-face exchanges, and over the phone four of them were still often addressed as women. Four of the speakers were willing to undergo vocal surgery to change their voices further and eleven rejected (any further) therapy believing it would not help. Fourteen of the subjects considered the voice change to be as important as the sex reassignment surgery. These results bring a more realistic view to the topic, and they strongly suggest that "becoming" or "being" a man may not be that simple and unproblematic, after all.

In the following chapters three acoustic phonetic studies provide both longitudinal case studies and cross-sectional accounts of transmen becoming men and performing maleness.

Chapter 3

The effects of androgen therapy on the biologically female voicing source

3.1 Overview

This chapter and the next (Chapter 4) present longitudinal acoustic analyses of transmen's voices during transitioning on testosterone. Although the effects of sex hormones on the human voice have been widely documented, very few studies address transgender individuals. Instead, studies have explored the vocal maturation of biological males (e.g., Titze, 1994), women's voices changing due to endocrine problems or the treatment of their symptoms (e.g., Pattie, Murdoch, Theodoros, & Forbes, 1998), and the ageing voice (e.g., Titze, 1994).

A small number of recent studies have begun to document the effects of exogenous androgen hormones in transmen (Adler & van Borsel, 2006; Constansis, 2008; Damrose, 2009; van Borsel et al., 2000, 2001). Although other researchers have presented pitch change data for transitioning transmen, the current longitudinal study is the largest to date, compared to only two subjects in van Borsel (2000). In addition, the current study is the first to analyse transmen's pitch changes using a psycho-acoustic pitch-scale.

Results of the current study do not support previously reported pitch range shrinkage (Adler & van Borsel, 2006; van Borsel et al., 2000), but indicate temporary shrinkage that virtually disappears by the end of the first year on testosterone. Second, the observed speaking pitches are significantly higher than the estimated optimal speaking pitches traditionally based on measured pitch floor values (e.g., Cooksey, 2000). Third, the current study finds previously undocumented patterns in which some transmen show a reverse J-pattern in pitch changes: after the initial drop and plateau, the habitual pitch slightly rises.

Several of these findings challenge existing theories of gender and transgenderism. Specifically, data from the current study do not support some theories of transitioning (e.g., Heath, 2006) that view the FTM transitioning process as striving for an unambiguously masculine end result. Similarly, the current results are not consistent with theories of gender that define gender in terms of a binary opposition of male and female categories, and transgenderism as the "switching" from one feature set to another. Instead, consistent with the Butlerian view, the data suggest that gender is continuous and transmen are often comfortable with a) moving in general towards a more masculine, as opposed to unambiguously masculine, self-representation, and b) keeping their vocal gender fluid and adjusting features of it to circumstances, which may even call for de-masculinizing vocal descriptors.

3.2 Background

3.2.1 Deriving the optimal voice fundamental frequency

In clinical practice it is assumed that in an ideal case a speaker's habitual pitch closely approximates their optimal pitch. By *habitual pitch* Case (2002) understands the modal or average pitch in a continuing sample of speech, the level around which normal pitch inflections occur. "Modal" here is used in both the phonetic and statistical sense of the word. It is the mode of the modal phonation range, that is, the pitch level used most frequently by the individual in typical conversational speech.

The most commonly used fundamental frequency elicitation procedures include eliciting free conversation, reading passages, and counting by the subjects. Zraick et al. (Zraick, Nelson, Montague, & Monoson, 2000; Zraick, Skaggs, & Montague, 2000) extensively document the lack of differences across speaking tasks, but report significant differences across speaking durations (Zraick, Birdwell, & Smith-Olinde, 2005) and social context (Zraick, Gentry, Smith-Olinde, & Gregg, 2006). There is also ample evidence that the habitual fundamental frequency is strongly influenced by the speaker's age, sex, race, language and the dialect within the language (Abu-Al-makarem & Petrosino, 2007; Braun, 1995; Traunmüller & Eriksson, 1995; Hollien, 1990; A. Hudson & Holbrook, 1982; Künzel, 1987, 1989; Lindh, 2006; Nolan, McDougall, de Jong, & Hudson, 2006; T. Hudson, de Jong, McDougall, Harrison, & Nolan, 2007; Pegoraro-Krook & Castro, 1994; van Bezooijen, 1995).

While the existence of a truly *optimal, optimum* or *best pitch* is often questioned, that concept is defined as a range of pitches that result in a slightly louder voice of clearer quality that can be produced with an economy of physical effort and energy and where the individual can phonate most efficiently (Boone, McFarlane, Berg, & Zraick, 2009). The question of habitual and optimal pitch is of special interest in transsexual and transgender populations. In the case of MTFs, the question is whether speech-language pathologists can safely retrain speakers to use a potentially very different range in their repertoire for everyday communication without voice quality deteriorating or becoming aphonic, losing intonational variety and projecting power, and without developing vocal pathologies. The little documentation available on the hormonal transitioning of FTMs usually shows substantial lowering of the fundamental frequency. There is no information, however, on whether speakers' habitual pitch lowers in tandem with the predicted or optimal speaking pitch. An additional concern is whether traditional optimal speaking fundamental frequency formulae are adequate to predict truly optimum speech performance for transgender speakers. Negotiating the distance between a "passing" pitch vs.a predicted or actual optimal speaking fundamental frequency is very likely a task for the speech-language pathologist.

Since the 1940s there has been substantial effort devoted to predicting or deriving the optimal or best pitch of speakers given their pitch floor and available pitch range. Based on Pronovost's (1942) and Fairbanks's (1960) so-called 25% Method, one school determines the optimal pitch to be 25% above the pitch floor, where the whole pitch range is treated as 100%. Britto and Doyle (1990) compared habitual (modal) pitch with derived optimal pitch values in a population of 40 young adults with no laryngeal pathology or vocal training. Their results indicate that habitual fundamental frequency in typical adults occurred much lower, namely, between 8% and 10% above the pitch floor.

It has been repeatedly observed that in everyday conversation, speakers use a relatively narrow range towards the lower end of their entire voice range. This is estimated to be around 3 ST above and below the average pitch for about 90% of the time for both short- and long-term sampling (R. Coleman & Markham, 1991). It has also been noted that eliciting the pitch ceiling yields less consistent trial-to-trial results than eliciting the pitch floor. Motivated by these two reasons, another school bases the derivation of the habitual pitch solely on the pitch floor. Previous research indicated that boys' speaking voices by the end of vocal maturation are 3-4 semitones above the lowest pitches in their vocal range (Cooksey, 2000; J. Killian, 1999; J. N. Killian & Wayman, 2010). Therefore, in this study each speaker's habitual speaking pitch will be identified and compared with the pitch floor.

3.2.2 The use of acoustic vs.auditory scales

Related to the accurate measurement of the acoustic correlate of pitch, the fundamental frequency, is the issue of the psycho-acoustic representation of the values. It is often convenient to measure absolute frequencies in Hz when pitch or change in pitch is studied. The pitch of a sound may be specified by the frequency of a pure tone whose pitch is judged to be the same as the pitch of that sound. But while the production of frequency is linear by nature, the perception of frequency is logarithmic. As a result, logarithmic or other non-linear auditory scales of frequency representation are required in models of auditory perception.

Historically, most psycho-acoustic scales went through repeated recalculations in order to approximate human perception, often taking age and task effects into consideration as well. As a result, the scales have multiple different formulae to convert from Hz values, and the output of the formulae can be markedly different. As an example, there are at least five formulae in use for Hz-to-Bark conversion: Liberman, Harris, Hoffman, and Griffith (1957), Tjomov and Gershuni (1971), Schroeder, Atal, and Hall (1979), Traunmüller (1990), and Zwicker and Terhardt (1980).

Out of the available psycho-acoustic scales, the musical semitone (ST) scale is a simple logarithmic transformation of the linear Hz scale:

$$ST = 12ln(f - value/f - reference)/ln(2) \tag{3.1}$$

The ratio pitch scale (Mel scale) and the Critical Band rate z (Bark scale) are linear below 500 Hz but logarithmic above:

$$e.g.,mel = 1127.01048ln(f/700 + 1) \tag{3.2}$$

$$e.g.,Bark = (26.81/(1 + 1960/f)) - 0.53) \tag{3.3}$$

The Equivalent Rectangular Bandwidth rate (ERB) scale is between linear and logarithmic below 500 Hz but logarithmic above 500 Hz:

$$e.g.,ERB = 11.17ln((f + 312)/(f + 14675)) + 43.0 \tag{3.4}$$

As is widely known, psychoacoustic scales such as these are important for adequately capturing the fact that a given pitch difference, say 100 Hz, is perceived as a different musical distance if it is a change from 400 Hz to 500 Hz than if it is a change from 100 Hz to 200 Hz. In order to illustrate the difference between the scales, in Table 3.1 on page 33 a change of 100 Hz is expressed in the most common psychoacoustic scales. The overview of non-linear auditory scales of frequency in Table 3.1 demonstrates that all non-linear scales accurately capture the fact that the difference between 100-200 Hz is perceived as a larger auditory distance than the difference between 400-500 Hz, all other factors being equal.

		Musical pitch	**ERB***	**Critical band† rate z**	**Ratio‡ pitch**
The distance	**100-200 Hz**	12 ST	2.47 ERB	1.18 Bark	132.74 Mel
between	**400-500 Hz**	3.9 ST	1.37 ERB	0.9 Bark	98.06 Mel

Table 3.1 : A comparison of methods to express the difference between two frequencies in Hz.

This phenomenon is recognized as particularly important in musicology and music-related clinical practice. For example, the approximate singing fundamental frequency ranges of four voice classes (as determined by the Voice Range Profile method (Titze, 1994) are introduced in Table 3.2 on page 34. The ranges presented here are the so-called concertizing ranges, which usually exclude the extreme one or two semitones of the full physiological range. As noticeable, the four lowest voice classes (bass, baritone, tenor and contralto) encompass a two-octave or 24-semitone range. When the distance between the pitch ceiling and the pitch floor are expressed in terms of

Hz difference, the perceptually equivalent ranges grow from 247 Hz (bass) to 440 Hz (contralto).

	min	**max**	**range** (Hz)	**range** (ST)	**mean**	**min to mean** (ST)
bass	82.4	329.6	247.2	24	98	3
baritone	98	392	294	24	123.5	4
tenor	130.8	523.3	392.5	24	164.8	4
contralto	146.8	587.3	440.5	24	174.6	3
mezzo-soprano	164.8	880	715.2	29	196	3
soprano	196	1174.7	978.7	31	246.9	4

Table 3.2 : Approximate range and speaking voice in the six major voice classes. Columns *min*, *max* and *mean* based on Tables 7.3 and 7.4 in Titze (1994, p. 187-188).

In order to maintain comparability in within- and between-subject comparisons, in this study the results measured by the author, as well as results obtained in Hz in previous studies, are presented in semitones (ST). The reason for choosing the ST scale over the other psycho-acoustic scales is two-fold. On one hand, being one of the few commonly used scales both in clinical practice and musicology, the results are easily interpretable for the widest possible audience. Additionally, the ST scale is the only scale with only one agreed-upon formula that has remained virtually unchanged in the past few centuries (Miller, 2002).

3.2.3 Hormones and their effects on the voicing source

It has been recognized that testosterone profoundly affects the development of the human larynx, both in males exposed to endogenous testosterone and chromosomally female speakers exposed to androgens, including female-to-male transsexuals and cisgender women with hormonal imbalances. The effect of androgens invariably manifests in the substantial and irreversible lowering of the fundamental frequency.

Life-span changes in the voicing source

During the life-span the changing morphology of the vocal fold tissues and larynx accounts for the bulk of the forces shaping the voicing source.

At infancy the membranous length of the vocal folds is approximately 2 mm. With an average growth rate of 0.4 mm / year for women and 0.7 mm / year for men, the adult membranous length is expected to be around 10 mm vs.16-17 mm, respectively (Titze, 1994, p. 179). At the same time, the cartilaginous length never really differs much, and differs only very slightly between adult men and women.

Between the ages of 10-18, there is a sudden increase in the rate of growth and laryngeal size of adolescents. It is generally agreed that girls' voices mature due to an enlargement of the larynx that is consistent with general body growth. Boys' speaking pitch, however, drops about an octave, and the membranous length of the vocal folds double due to the effect of testosterone (Greene & Mathieson, 1989, p. 66). At the same time, the bulk of the thyroarytenoid muscle, as well as the vocal fold thickness and mass, show an increase in the male pubertal larynx (Titze, 1994, p. 181–182).

After the changes in puberty, the vocal source is fairly stable up into the 60s (cf. Figure 7.10. in Titze (1994, p. 183). When summarizing general trends, Titze (1994, p. 184) notes that the female habitual pitch changes more unidirectionally than the male pitch does. With this he calls attention to changes in the habitual pitch after the mid-age plateau between the ages 20 and 50 years of age. While with the diminishing menopausal estrogen levels women's pitch tends to drop, there is evidence that the diminishing testosterone levels cause a slight pitch rise in males (Hollien, 1987). The reason for that rise are hypothesized to be due to the general loss of muscle tissue. In the literature no other account of rising male pitch was found.

Female endocrine vocal changes: Hyperandrogenism

After the predictable changes during the lifespan, in this section I discuss a medical condition, hyperandrogenism, that bears upon the issue of the effects of male hormones on the chromosomally female voicing source. Hyperandrogenism is a medical condition characterized by the excessive production of male hormones by either sex.

The principal mechanisms resulting in hyperandrogenism are either an increase in circulating androgen levels or an intrinsic androgenic activity of a particular drug. The natural causes of hyperandrogenism may be due to adrenal tumours, ovarian tumours or 21-Hydroxylase deficiency. The drugs most commonly responsible for the virilizing effects are anabolic steroids, progestins and antiepileptics, as well as the use or abuse of testosterone, androstenodione and dehydroepiandrosterone (DHEA) in athletic performance or treating the symptoms of menopause, chronic fatigue, hyposexuality and endometriosis (Neraud & Dewailly, 2010, p. 121).

The symptoms of hyperandrogenism are a combination of weight gain, acne, hirsutism, breast atrophy, increased musculature, clitoromegaly, greasy skin, headaches, depression and deepening voice. Contrary to popular belief, the presence of such virilization is not a feature of Polycystic Ovary Syndrome (PCOS) (which has a higher than average incidence rate among lesbians and transmen (Meier, in-press), but instead only occurs with more severe hyperandrogenism (Guzick, 1998).

Baker (1999) and Boothroyd and Lepre (1990) hypothesize that the virilization of the voicing source during hyperandrogenism is caused by the vocal folds getting inflamed or the muscle tissue altering and thereby bringing on altered aerodynamic properties. Gerritsma, Brocaar, Hakkesteegt, and Birkenhger (1994) observed increased vocal fold stiffness which they explained with increased local muscle mass with concomitant loss in connective tissue.

Importantly, case studies of hyperandrogenism establish that this noticeable deepening of the habitual pitch happens in the first three months of androgen therapy, and is undesired, yet permanent and irreversible even on cessation of the androgen therapy (Baker, 1999; Barbieri, Evans, & Kistner, 1982; Boothroyd & Lepre, 1990; Mercaitis, Peaper, & Schwartz, 1985; Newman & Forbes, 1993; Pattie et al., 1998; Wardle, Whitehead, & Mills, 1983).

In sum, androgens not only lower the pitch in both genders, but they do so permanently. The change in pitch is due to the increase of vocal fold mass with possible contribution from vocal fold stiffness as well. The irreversibility of the changes has been exemplified with cismales, cisfemales with hyperandrogenism, and male-to-female transsexuals.

3.2.4 Previous work on the transitioning FTM voice

van Borsel et al. (2000) and Adler and van Borsel (2006) present longitudinal data of (the same) two FTMs, with data points collected about every two months. The subjects performed phonatory range exercises, sustained the vowel /a/ and read the

North Wind and the Sun passage. No spontaneous speech was analysed and the elicitation of sustained vowels was analysed and discussed only in terms of short-term perturbations. Fundamental frequency measurements were extracted from both vowels and passages, averaged and plotted against time. Subject S1 was 22;4 years old at the commencement of data collection and his hobbies included singing which he practiced about an hour a day. S2 was 37;10 years old, a smoker (a pack of cigarettes a day) without any known voice training.

Damrose's (2009) subject was a 33-year-old semi-professional jazz singer. He was treated with testosterone ethanate 200 mg intramuscularly bimonthly. A sustained /a/ at habitual loudness and pitch was recorded for 5 seconds once a month for 20 months. At the same time laryngostroboscopy was performed using a rigid endoscope and digital stroboscopy system. Mean fundamental frequency, shimmer, noise to harmonics ratio and voice turbulence index (VTI) were measured in his production. While the methods do not include information about if or how other speech materials, such as phonatory range exercises, reading passages or spontaneous speech, were elicited, changes in maximal and minimal mean fundamental frequencies over time are reported too.

The changes documented in van Borsel et al. (2000), Adler and van Borsel (2006) (labeled as AvB 1 and 2), and Damrose (2009) (labeled as D 1) are summarized in Table 3.3 on page 39. The results in these studies were originally reported in Hz and are shown as ST according to formula 3.1 on page 32.

The table allows examination of the fall in habitual fundamental frequency during

Subject	Time	Mean F_0 (Hz)	Min F_0 (Hz)	Max F_0 (Hz)	Min to max (ST)	Min to mean (ST)
D 1	pre	228.47	140.26	338.49	15.2	8.4
D 1	16 mo	112.74	90.75	201.07	13.7	3.7
change (ST)		-12.2	-7.53	-9	-1.5	4.7
AvB 1	pre	215-21	165	800	27.3	4.6
AvB 1	13 mo	155	105	500	27	6.7
change (ST)		-5.6	-7.8	-8.1	0.3	-2.1
AvB 2	pre	160	125	525	24.8	4.3
AvB 2	12 mo	132	82	350	25.1	8.2
change (ST)		-3.3	-7.3	-7	0.3	-3.9

Table 3.3 : F_0 change of subjects in the two previously published studies.

the first year on testosterone. Additionally, it is clear that both the highest and lowest pitches in the production of each subject fell steadily. The mean pitch decreased by 3 to 12 ST, the pitch floor by about 7 ST, and the pitch ceiling by 7 to 9 ST.

Whereas the original authors argued, based on the Hz-based range values, that transmen's available pitch range radically narrowed as a result of the testosterone therapy, this conclusion is unsupported after conversion to a ST scale (c.f. Column 6 "Min to max"). Adler and van Borsel (2006, p. 144) repeat verbatim van Borsel et al.'s (2000, p. 434) observation "Clearly, the pitch range of S1 was seriously reduced as a result of the hormone therapy. This reduction appears to be the result of a loss in the high tones which is not fully compensated for by a gain in the lower frequencies." Similarly, Damrose (2009, p. 111) concludes that "Overall there was a marked shift and contraction of pitch range." As a result, Adler and van Borsel (2006, p. 144) remind clinicians that "Clients who do sing, and certainly professional singers, should be warned that their pitch range will irreversibly alter consequent to the administration of the cross-gender hormones." However, as the values in column 6 demonstrate, the net change over 12 months was in the range of one semitone for all three subjects, a musically insignificant change that is well within the margin of error for measuring pitch across multiple measurement sessions (i.e., a trivial trial-to-trial variation within speakers).

An additional problem with interpreting Damrose's results lies in the elicited range. As a clinical rule of thumb, a range of 24+ ST (2+ octave) can be expected even from speakers who are not musically trained. Damrose's subject, a trained, professional and active pop and jazz singer, showed a pre-testosterone range of only 15 ST and post-transition range of only 14 ST. A professional singer would be expected to possess a range of at least 30 ST, if not larger. It is not clear whether the range restriction is due to an atypical elicitation technique or subject idiosyncrasy.

Finally, the results from the three studies in the last column indicate that for some subjects the habitual pitch increasingly elevates above the pitch floor, expanding the distance to as much as 8 ST.

3.3 Hypotheses to be tested

To test predictions growing out of the literature just discussed, a longitudinal study of six transmen on testosterone was conducted. As measures of the fundamental frequency provide information about how speakers use their voice, maximum phonational frequency range (MPFR), mean habitual pitch and the distance between the pitch floor and mean habitual pitch are documented in transitioning transmen.

Hypothesis 1. *Based on previous literature on habitual and optimal speaking pitch (e.g., Cooksey, 2000), when a transitioning transman achieves a cismale-typical fundamental frequency range with a pitch floor of 80-100 Hz, he is predicted to have an optimal pitch range around 100-126 Hz.*

Hypothesis 2. *Based on the clinical warnings of van Borsel et al. (2000), Adler and van Borsel (2006) and Damrose (2009), the pitch range of transmen is expected to go through a narrowing phase during the first year, the duration of which is unknown.*

Hypothesis 3. *Based on the irreversible quality of androgen-induced changes (e.g., Pattie et al., 1998), transmen are expected to show only downward movement of pitch measures. Upward movement is considered impossible on physiological bases.*

3.4 Subjects

The speech of seven non-smoking, musically untrained female-to-male transgender individuals (mean age 32.3, min = 23, max = 47, c.f. Table 3.4) on page 42 undergoing androgen therapy was documented about every two weeks over a 12 month period during their hormonal transitioning from female to male. For five subjects, recordings were made approximately every two weeks. One out-of-town subject's (JS, 38 years old) production was documented only before the commencement of testosterone therapy and one year into it. One subject, NL, moved out of town after about 200 days after the commencement of the testosterone therapy, but during that time multiple measurements were made of his production. Subject EX self-documented his transitioning with weekly-biweekly recordings. His data were used in tracking the changes in the mean fundamental frequency, but the nature of the recorded data did not allow for pitch range measurements.

In weekly instalments the speakers self-administered 75-250+ mg testosterone cypionate intramuscularly. However, subjects did not always use exact amounts and their dosage was often re-titrated according to their needs. Therefore the cited dosage should be viewed as indicating the intended dosage and not as factual information on the subjects' actual testosterone level.

Subject ID	**age (yr)**	**available data**	**measurements**
CM	23	3 years	mean, min, max F_0, F1-F3
MS	23	1 year	mean, min, max F_0, F1-F3
EX	31	5 years	mean F_0
TJ	32	1 year	mean, min, max F_0, F1-F3
LW	38	1 year	mean, min, max F_0, F1-F3
JS	38	before & 1 year into T	F1-F3
SB	30	before & 1 year into T	F1-F3
NL	47	200 days	mean, min, max F_0
average	32.3		

Table 3.4 : Subjects in the longitudinal studies, with ages at the commencement of the hormone therapy and available data as of January 2011.

3.5 Materials and procedures

Subjects provided both spontaneous and read speech samples (the Rainbow Passage, Houston Urban English Survey Passage (Niedzielski & Koops, 2011), and Comma Gets a Cure). They also performed pitch range exercises that mapped the extremes of their available vocal range. The recordings were made with a Shure SM10A head-mounted unidirectional dynamic microphone onto a Marantz PMD730 solid-state recorder at a 44.1 kHz/16 bit sampling rate. In this study only the read speech samples and the pitch range exercises are analysed.

Reading passages were used because of their documented within-speaker consistency of pitch as evidenced by Ladd's work (1997, p. 64–66). His production experiments, in which speakers read controlled texts, consistently revealed regularities of the fundamental frequency at particular points in utterances, including the low pitch at the end of a sentence-final fall, and, to lesser extent, the height of accentual peaks, which can vary strongly with emphasis, place in utterances, etc. In general, though, when other factors are carefully controlled, using a reading passage set is likely to elicit a certain degree of within-speaker range stability. Subjects were asked to read the short passages once per session as if they were reading it for a friend.

Maximum phonational frequency range (MPFR) (also referred to as pitch range, vocal range, phonational frequency range, physiologic frequency range of phonation, and musical frequency range of phonation) is an indirect measure of a subject's laryngeal function during voice production that typically encompasses the range of frequencies from highest to lowest. In addition to reflecting an individual's vocal ability, it is also often used to indirectly assess the physical condition of the phonatory mechanism.

The pitch range exercises employed were tailored to the subjects' musicality. If the subject knew vocal warm-up techniques, they were asked to (discrete-step) scale or glide as low and as high as they physically could. For other subjects various imageries were created, such as "Imagine you are on a roller coaster. Pretend that you're following the roller coaster with your voice," or "Imagine you are imitating a siren." Some subjects found deliberate yawning and sighing conducive to reaching their lowest pitch, and performing the Texas A&M "whoop" a good way to reach their highest. Subjects were asked to repeatedly produce vocal range maxima until three successful tokens of the lowest and highest pitch values were collected. The three tokens deemed to be delivered at maximum performance were then averaged. The three averaged values were within 40 Hz from each other at the high end and about 2 Hz at the low end. As a reviewer pointed out, maximum performance tasks are not averaged in clinical practice. Instead, the mathematical maximum is used in characterizing pitch ranges. In further work on FTM pitch ranges, I am planning to recalculate the results in a way that allows clinical comparisons.

The recordings were carried out prior to the onset of hormone therapy and fortnightly thereafter up to 12 months. If the subject agreed and was available, further recording sessions were conducted past the one year mark. If a subject relocated, his data until relocation were used even if he was documented for a period shorter than a year.

All but one subject made the recordings in a sound-treated booth at Rice University. Subject EX made his recordings on a personal computer with a series of simple external microphones over the years. However, because pitch and formant measurements are very robust even against considerable background noise and the self-documentation had exceptional temporal resolution, EX's data were included and analysed in this study.

3.6 Analysis

The second sentence of the Rainbow Passage ("*The rainbow is a division of white light into many beautiful colours.*") was extracted from the whole recorded passage. The final two syllables were removed from the analysis as they often contain vocal fry (creaky portions) which may skew the analysis. In Praat, pitch objects were created with a F_0 tracker set to a 70-350 Hz range using the cross-correlation method. The pitch contours were then manually corrected for octave and fifth jumps, as well as whistled and otherwise mistracked fricatives. Finally, the minimum (also referred to as pitch floor), mean and maximum (a.k.a. pitch ceiling) fundamental frequency values were determined.

The physiological pitch range, that is, the distance between the pitch minimum and maximum regardless of quality or aesthetics, was expressed in semitones using the Hz to ST conversion formula presented in Section 3.2.2. The distance between the reading sentential mean and the pitch floor was also calculated and expressed in semitones. Finally, the reading sentential means, minima and maxima were plotted against the time on testosterone for each subject. In order to highlight visual trends, a LOWESS (Locally WEeighted Scatterplot Smoothing) function with a smoothing span of 0.5-0.8 was fitted over the data points in R.

3.7 Results

An overview of the changes of the pitch descriptors during the first year of testosterone therapy are listed in Table 3.5 on page 47. The first column indicates the mean fundamental frequency in the second sentence of the Rainbow Passage, the second and third the extreme fundamental frequency values measured in the range exercises. The total physiologically available pitch range in ST can be found in the column *Min to max range*. The final column shows the distance between the pitch floor and the mean reading pitch in ST.

The results in Table 3.5 show that the sentential mean fundamental frequency decreased substantially in some subjects. On average the one-year pitch means are about 3-11 ST lower and the pitch minima about 6-12 ST lower than the pre-testosterone values. Pitch maxima changed differently among subjects. CM and NL lost about 8-9 ST from their upper range, while JS, LW, MS and TJ gained 4-10 ST in the upper range by the end of their first year on testosterone.

As demonstrated in the fifth column in Table 3.5 subjects CM and NL experienced a slight decline in pitch range, while the others experienced that their pitch range widened over the first year on testosterone. The location of the mean pitch value within the pitch range also changed. With the exception of JS and MS, subjects' mean pitch before transitioning was higher than the predictable 3-4 ST above the pitch floor. This distance grew in the case of every subject in this study. In the case of subjects LW and TJ this means that they spoke a full octave (12 ST) above their pitch floor.

In the following sections the summarized descriptors are considered in turn.

Subject	**Mean F_0** (Hz)	**Min F_0** (Hz)	**Max F_0** (Hz)	**Min to max range** (ST)	**Min to mean range** (ST)
CM	175.9	117	773.5	32.7	7.06
CM	135.9	79.3	450	30.05	9.33
change (ST)	4.47	6.73	9.38	2.65	-2.27
JS	192.5	147.4	622.1	24.93	4.62
JS	101.5	71.9	734.7	40.24	5.97
change (ST)	11.08	12.43	-2.88	-8.29	-1.99
LW	165.2	108.1	414.4	23.26	7.34
LW	137.3	67.7	532.8	36.24	12.24
change (ST)	3.2	8.1	-4.35	-12.98	-4.9
MS	202.3	177.4	426.7	15.19	2.27
MS	138.9	117	685	30.5	2.97
change (ST)	6.51	7.21	-8.19	-15.31	-0.7
NL	176.8	128.6	907.2	33.82	5.51
NL	126.7	82.3	465	29.98	7.47
change (ST)	5.77	7.73	11.57	3.84	-1.96
TJ	199.6	136.4	571.3	24.8	6.59
TJ	173.1	80.5	781.8	39.36	13.25
change (ST)	2.47	9.13	-5.43	-14.56	-6.66

Table 3.5 : F_0 change over 12 months on testosterone (or in the case of NL, 186 days)

3.7.1 Mean fundamental frequency

Besides calculating the amount of the changes after one year on testosterone, it is important to characterize the timing of the changes. Some of the most frequent questions from FTMs commencing testosterone therapy address how soon the changes are noticeable and how long it takes for the mean fundamental frequency to reach its lowest and settle. In order to illustrate the changes in the six subjects, their individual fundamental frequency curve was plotted as a function of time on testosterone.

Subject CM

When considered on the linear Hz scale, the fundamental frequency of CM seems to have lowered in three major phases, the second commencing after about 2 months on testosterone and the third around the 7th month (cf. Figure 3.1 on page 49). The changes slow down when he reached around 140 Hz after approximately one year on testosterone. After about 600 days on testosterone, the fundamental frequency slowly started to rise from around 123 Hz and re-assumed values already reached by the end of the first year. Currently CM's mean fundamental frequency is around 142 Hz, which is 3 ST above his lowest four fundamental frequency measurements of 123 Hz. This pattern of change of fundamental frequency is going to be referred to as the reverse J-pattern in the rest of the study.

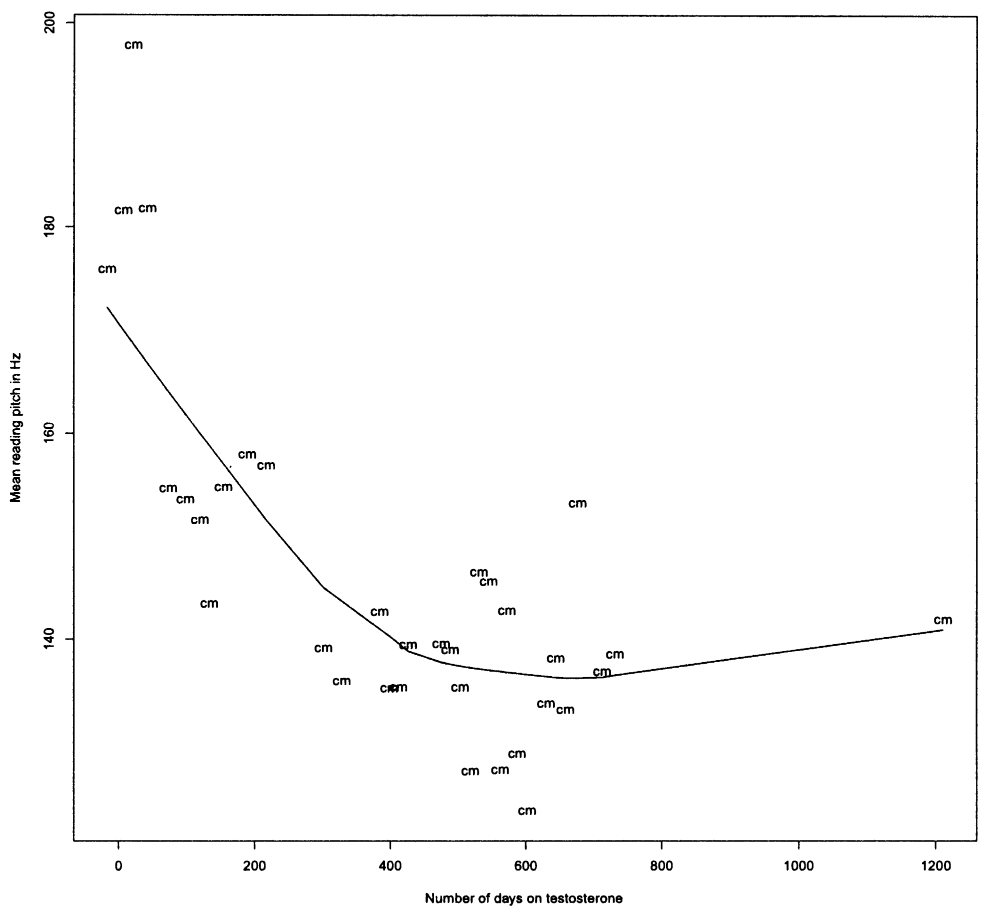

Figure 3.1 : Subject CM's mean reading fundamental frequency changes as a function of time on testosterone.

Subject EX

For EX, the first drop of 100 Hz (which in his case was about an octave or 12 ST) happened in a largely linear fashion during the first 200 days of androgen therapy (cf. Figure 3.2). Following this, the next 20 Hz of lowering happened in the next 7-8 months, when he reached his lowest mean fundamental frequency, an F_0 of about 100 Hz. As seen in CM's production of the reverse J-pattern above, after the first 400 days, EX's mean fundamental frequency started slowly but steadily to rise. Currently his reading fundamental frequency is around 115 Hz, about 2-3 ST above what it was at his lowest. However, an important difference between CM's and EX's production needs to be pointed out. While the rising mean fundamental frequency was borne out of a large fluctuation in the case of CM, in EX's production the mean fundamental frequency rose from a settled, barely varying pitch profile.

Subject LW

Similar to subject CM, LW also shows a slight stepwise change, with minor breaks after 150 days and 300 days on testosterone (cf. Figure 3.3). After about 10 months, his mean reading fundamental frequency seems to center around 130 Hz, but with large session-to-session fluctuations. Even though LW agreed to participate only for a year, he provided further data after his oophorectomy (ovary removal) around 500 days into his hormonal transitioning. About 2-3 months after the surgery, which eliminated his estrogen production, and after which LW's testosterone intake was halved, LW's mean fundamental frequency dropped to 120 Hz. This fundamental

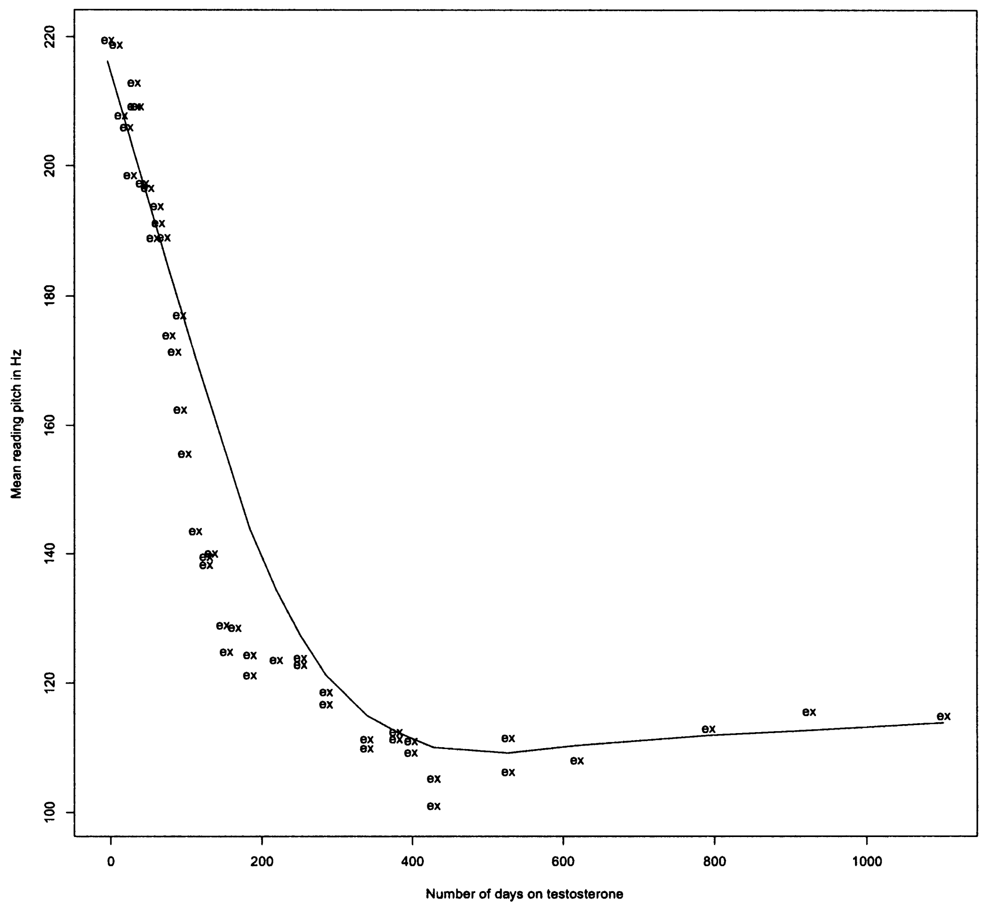

Figure 3.2 : Subject EX's mean reading fundamental frequency changes as a function of time on testosterone.

frequency value (120 Hz) was further corroborated by the measurement taken a further 200 days after that. While the data resolution is not optimal in the last 400 days of the sampling, it indicates that the oophorectomy exerted an influence on his mean sentential fundamental frequency.

Subject MS

The mean sentential fundamental frequency of MS decreased very sharply in the first 2 months, during which he completed the bulk of his vocal transitioning (see Figure 3.4). In the following 5 months his fundamental frequency declined further but at a slower pace. MS relocated after 200 days of participation in the study, so there are no data between days 200 and about 400. Due to this, it is not possible to ascertain if his changes stabilized around 120 Hz or if his fundamental frequency continued to lower.

However, about 400 days into his hormonal transitioning, he agreed to be recorded during a visit to Houston. That measurement was substantially higher than the previous measurement. MS's data are thus open to at least two interpretations. On one hand his pitch might have settled around 130-140 Hz with large fluctuations, like the "rise from chaos" pattern of CM. On the other hand it is possible that his pitch also started moving back up after a period of settling at a lower pitch, as in the case of EX whose pitch showed more of a "rise from stability" picture.

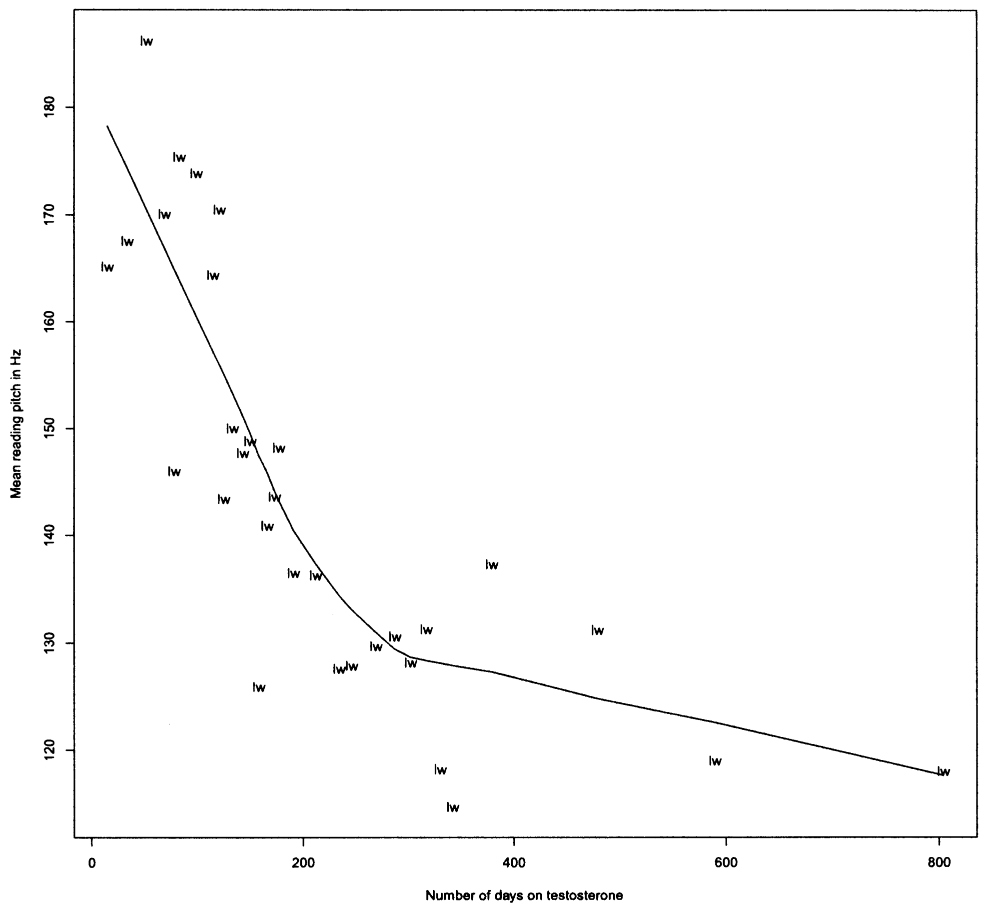

Figure 3.3 : Subject LW's mean reading fundamental frequency changes as a function of time on testosterone.

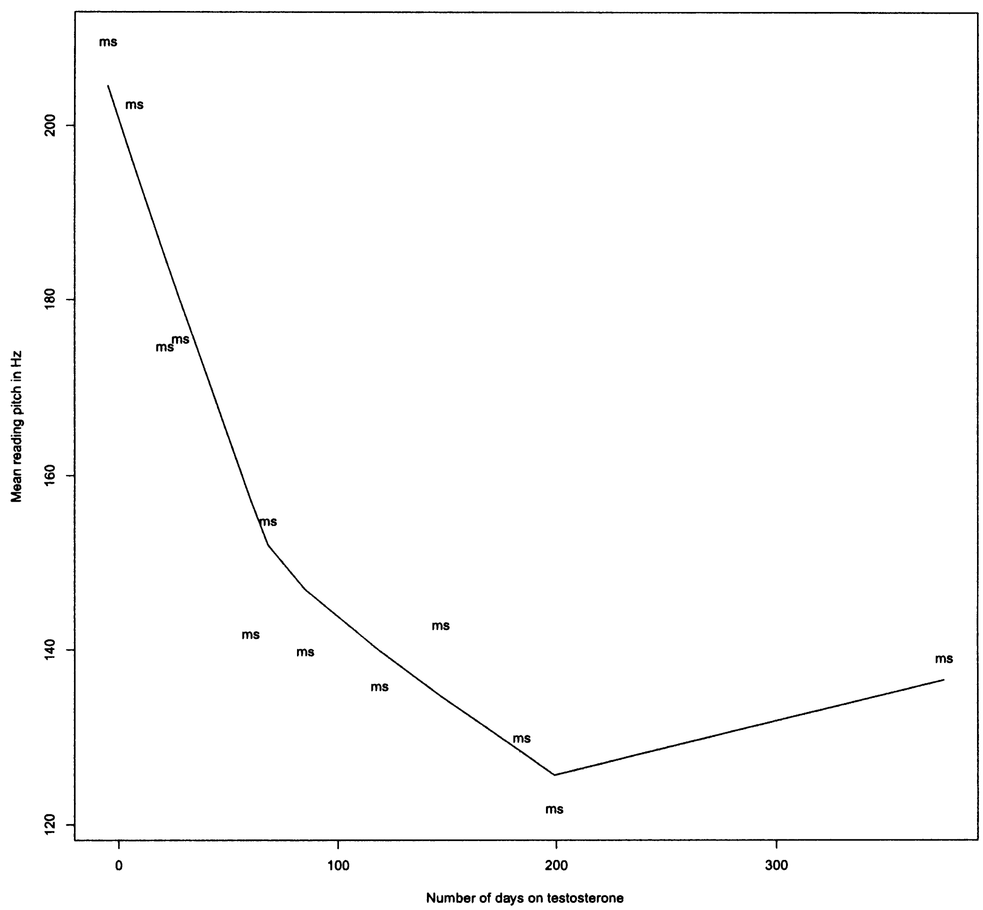

Figure 3.4 : Subject MS's mean reading fundamental frequency changes as a function of time on testosterone.

Subject NL

NL started his hormone therapy after menopause, so it was expected that testosterone would have a profound and fast effect on his pitch. NL's initial mean reading pitch of about 180 Hz fell to about 130 Hz within the first 50 days after the commencement of androgen therapy (cf. Figure 3.5). During the next 150 days, before NL relocated and dropped out of the study, his pitch did not lower further.

Subject TJ

For family reasons, TJ chose to transition slowly. He therefore took about half the amount of testosterone the other FTMs in the study were prescribed (50-75 mg testosterone cypionate instead of the usual 100-150 mg). It is noticeable that his vocal transitioning is very gradual (cf. Figure 3.6), and fairly linear in nature. Over the 12 months, his pitch slowly lowered from about 200 Hz to 170 Hz. TJ stopped taking testosterone after one year for personal reasons.

Summary of overall pitch changes

Anecdotal evidence from the six subjects above seems to show that the speed, but not necessarily the magnitude, of the F_0 drop is proportional to the testosterone-estrogen ratio. Subjects who took larger amounts of testosterone, or whose estrogen production was (virtually) non-existent (cf. the post-menopausal and post-oophorectomy subjects) demonstrated a faster, more hyperbolic change. In order to help the reader compare the subjects, every subject's pitch change was re-plotted in Figure 3.7 on

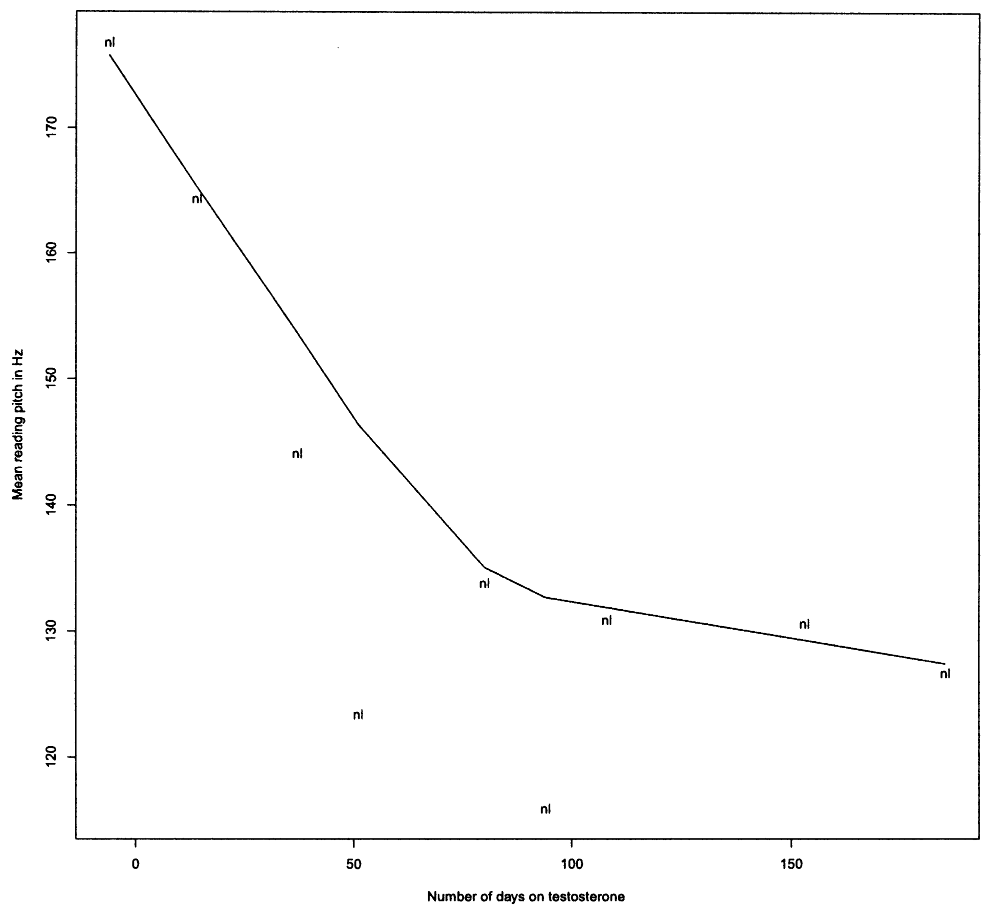

Figure 3.5 : Subject NL's mean reading fundamental frequency changes as a function of time on testosterone.

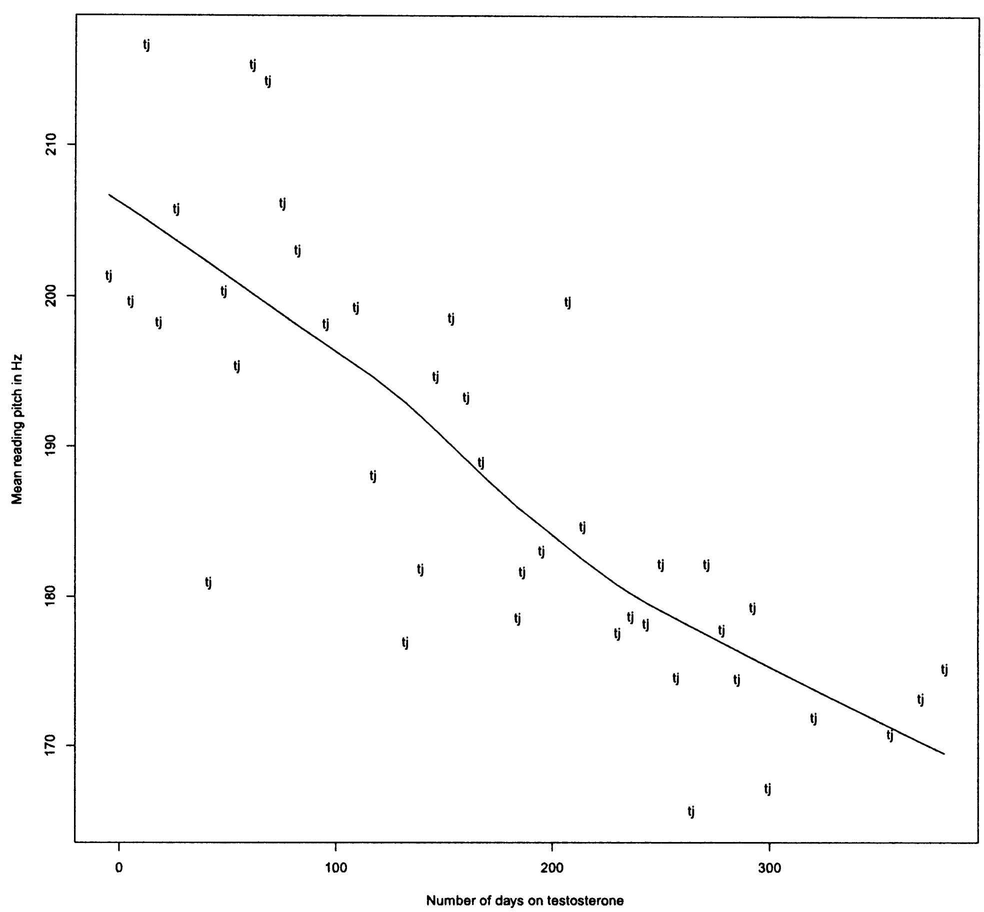

Figure 3.6 : Subject TJ's mean reading fundamental frequency changes as a function of time on testosterone.

page 59.

3.7.2 Pitch floor to mean fundamental frequency

Figure 3.8 on page 60 shows that the distance between the pitch floor and the fundamental frequency mean increases as a function of time on testosterone for four speakers in the first 100 days. For one speaker, TJ, this increase persists through the entire first year. As a result, by the end of the first year TJ spoke over an octave (13-14 ST) higher than his pitch floor, a fact that belies the restricted success of his mean habitual fundamental frequency lowering as discussed in the previous section.

For LW, who spoke higher in his range, the distance between the pitch mean and pitch floor slowly diminished from 12 to about 7 ST.

After an initial narrowing in the first 200 days, CM's floor-to-mean range showed steady increase. This increase persisted for the rest of the time CM participated in the experiment, that is, another 1000 days.

MS showed remarkable consistency in speaking about 2-3 ST above his pitch floor regardless of time in transitioning.

3.7.3 Vocal range

The absolute size of vocal range, i.e. the distance between the physiological pitch floor and pitch ceiling, was calculated for each recording session and plotted as a function of time in ST. Figure 3.9 (on page 62) allows the examination of the similar behaviour of subjects CM and LW. These two subjects' productions reached their smallest (i.e., the narrowest available frequency range) by around 200 days into their androgen

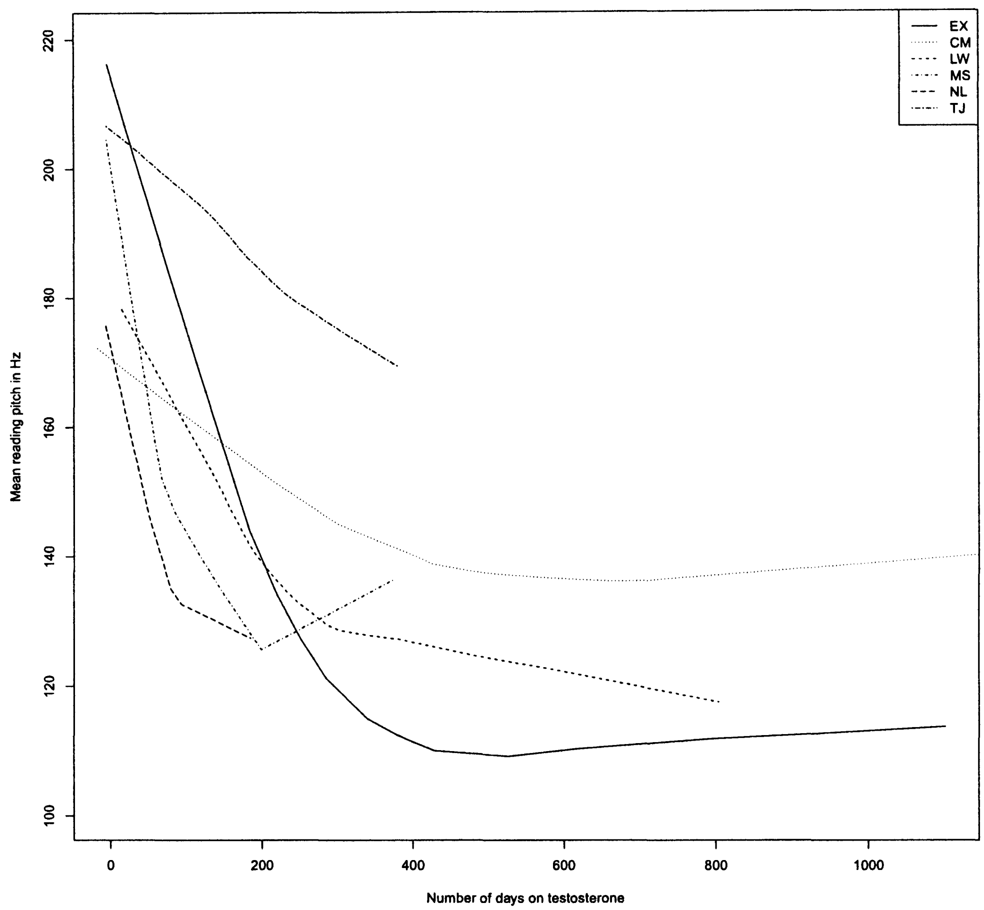

Figure 3.7 : Mean reading fundamental frequency changes as a function of time on testosterone.

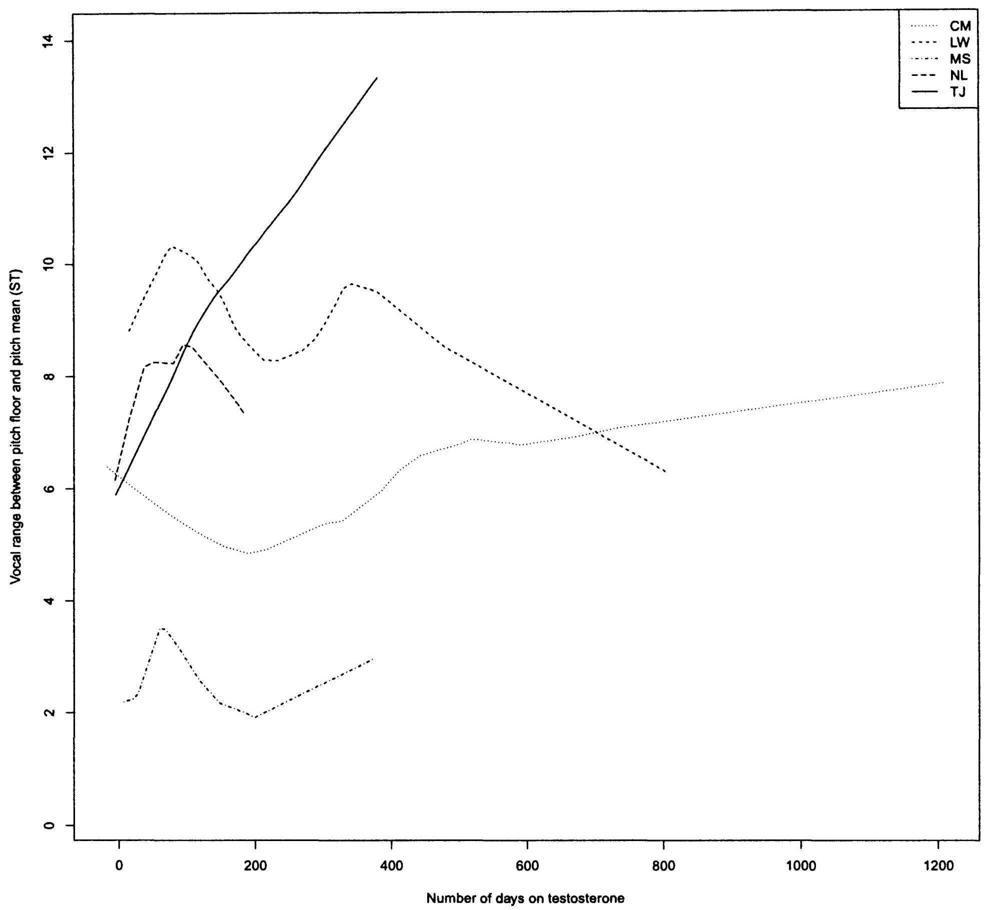

Figure 3.8 : Distance between pitch floor and mean reading pitch as a function of time on testosterone.

therapy. After a short lull, both subjects' range expanded and eventually surpassed their original range. NL's production is somewhat similar in the respect that it shows a strong decline in the range by the time he reached 200 days in the experiment. MS and TJ showed a rising pattern all through their time in the experiment.

When comparing Figure 3.9 with Figure 3.10 (on page 63) tracking the pitch ceiling and floor, it becomes clear that the observed reduction in the vocal range is because of the loss of high frequencies during transitioning. This loss is noticeable in CM's, LW's and NL's production. Arguably, the slowing down or "break" in the increase in the pitch ceiling in the production of MS and TJ around the same time period is a manifestation, albeit in a milder form, of the same effect.

3.8 Discussion

This study aimed to collect data on the transitioning fundamental frequency of FTMs and place it in the context of existing data in the literature.

Based on previous literature on habitual and optimal speaking pitch (e.g., Cooksey, 2000), it was expected that when a transitioning FTM achieves a cismale-typical fundamental frequency range with a pitch floor of 80-100 Hz, he will have a habitual speaking fundamental frequency of about 100-126 Hz, or about 3-4 semitones above the pitch floor.

Although all transmen in the experiment reached a pitch floor of 80-100 Hz, they did not speak at their predicted speaking fundamental frequency, but up to an octave (6-13 ST) higher. Since both the pitch floor and the pitch ceiling changed in all cases,

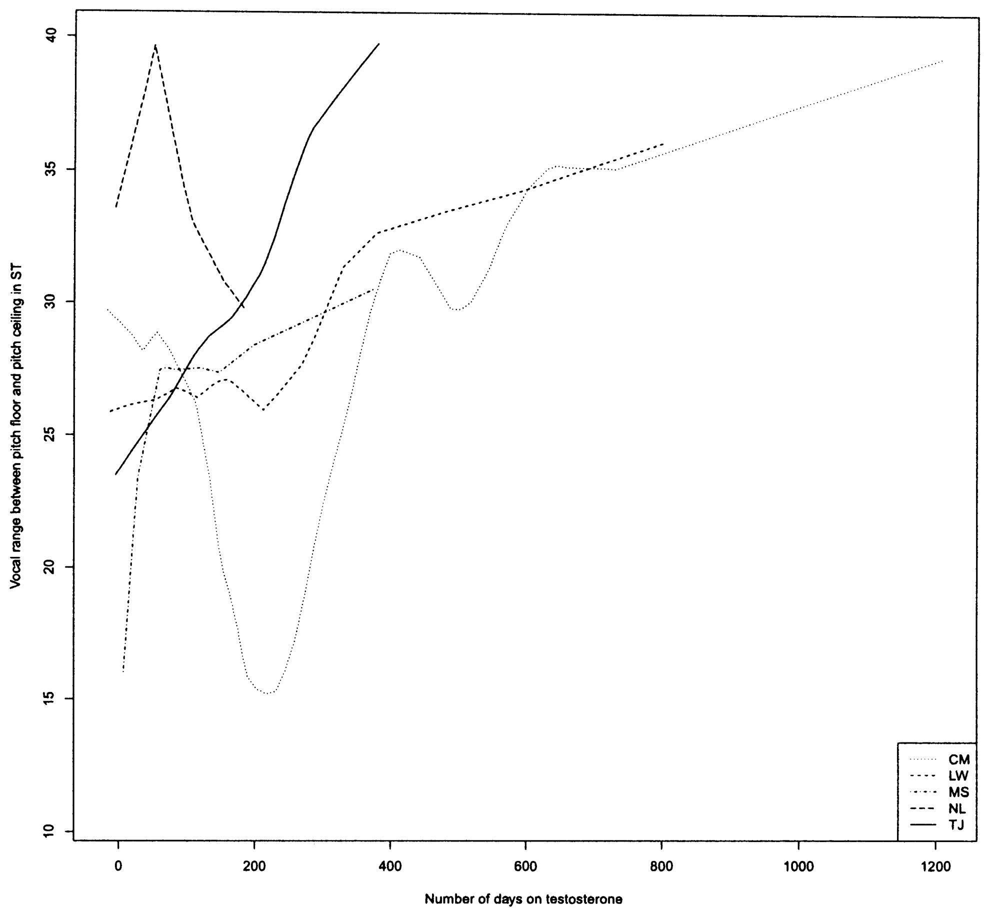

Figure 3.9 : Available vocal range as a function of time on testosterone.

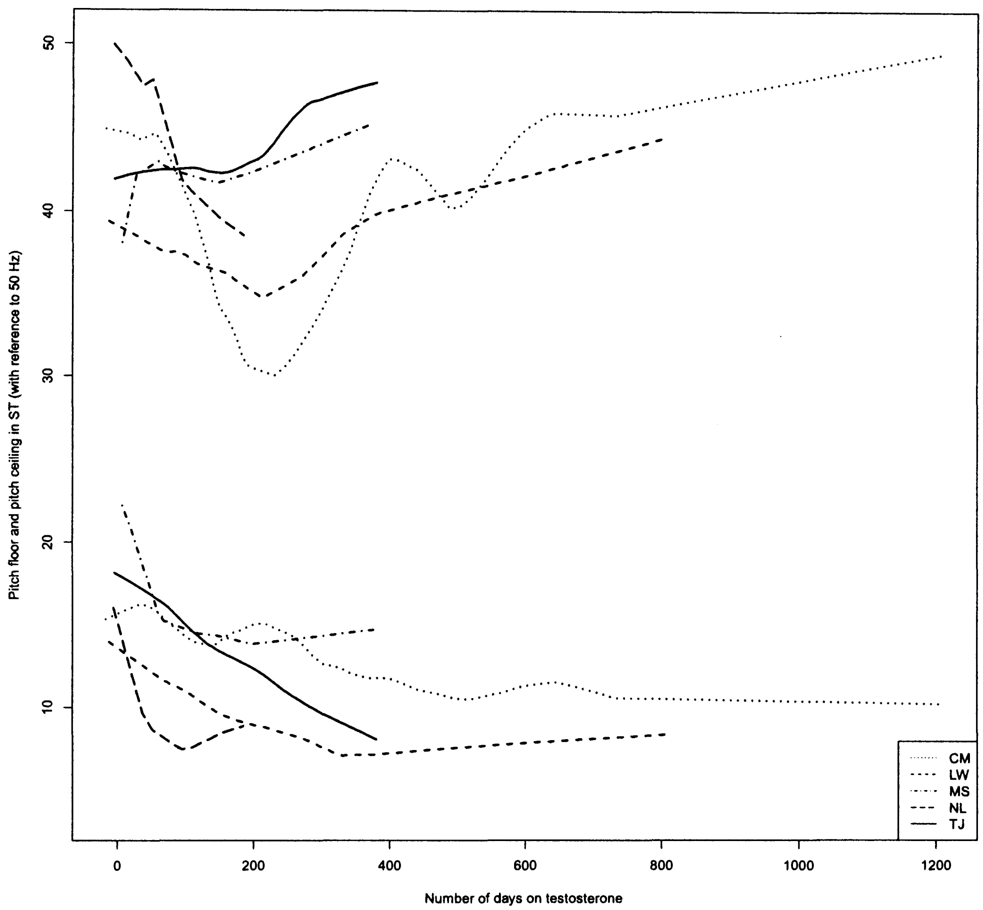

Figure 3.10 : Changes in the voice pitch and ceiling as a function of time on testosterone.

this indicates that FTMs' speaking voices are not situated as low in their range as they used to be before transitioning. This in turn means that their physiological range expands and lowers more than their habitual range does. The reason for the unexpectedly large pitch floor to fundamental frequency mean distance within the range can be due to life events, such as a change of sex or gender identity, a change in the speaker's need to be able to pass as a male member of the society or the perceived contribution of voice vs.physical changes during transitioning.

Another possible reason is that the transmen may have actively or passively resisted speaking in a lower (chest) voice. In LW's account, when he felt a strange "rumbling" in his chest he usually re-calibrated himself to a higher pitch, unless it was a formal situation demanding authoritative behavior, in which he maintained the low pitch. As another example, TJ would often ask whether his voice had changed. By that he meant an observable change in the mean fundamental frequency which he himself did not perceive as changing. He was indeed correct: over one year on testosterone his mean habitual pitch changed from 200 to 170 Hz, which is a change of 3 ST. At the same time, however, his pitch floor lowered from 130 to 80 Hz, or over 8 ST. If his pitch had recalibrated to a predicted 4 ST over his pitch floor he would have been speaking at around 100 Hz, which is an unambiguously masculine pitch. When asked if he tried to play with his voice and speak lower, TJ's answer was that he tried and "It sounded weird."

LW, CM and TJ all had problems passing well into their time on testosterone when their predicted optimal pitch, but not their habitual pitch, was unambiguously masculine. These results make a strong case for including speech-language pathologists (SLP) in the routine medical care of transmen. The role of an SLP would include both mapping the habitual speaking fundamental frequency within the available pitch range of transmen, and facilitating behavioural changes or making referrals to other professionals, such as endocrinologists and ear-nose-throat specialists (ENTs). SLPs are able to make important distinctions between transmen reporting problems relating to the habitual speaking fundamental frequency.

One type of problem is when transmen exhibit little or no lowering of the habitual fundamental frequency and the pitch maxima, even though they have been on testosterone for over a year. This is a physiological-hormonal problem, which needs to be investigated with the help of an SLP, an ENT and / or an endocrinologist. The other kind of problem would be reported by individuals who do have access to cismale-normative fundamental frequencies (which they may or may not know about) but do not seem to be able to speak at those frequencies habitually. These transmen would then need to be re-trained to realize the existence of and then adjust to frequency ranges that are conducive to passing or presenting as the desired sex or gender. Naturally, SLPs would not need to work with the third group of individuals who, while having access to cismale ranges, exhibit higher-than-cismale habitual fundamental frequency values *and* are happy with the way they are read (be it gay, queer, androgynous, effeminate, or, on occasion, female).

Another hypothesis addressed was that the pitch range of transmen was expected to permanently narrow because of the loss of high frequencies (van Borsel et al.

(2000), Adler and van Borsel (Adler & van Borsel, 2006; Damrose, 2009; van Borsel et al., 2000). This hypothesis was tested in two different ways in the present study.

First, a meta-analysis was carried out on the previously existing data (Adler & van Borsel, 2006; Damrose, 2009; van Borsel et al., 2000). In the analysis the results published in acoustically relevant Hz values were recalculated in the perceptually driven semitone scale. This allowed more accurate comparison both within and between individuals, which was not possible previously. While the clinical warnings in the above three studies appeared warranted based on the ranges measured on the linear Hz scale, converted to ST their subjects showed no change of available range as a result of the testosterone treatment. The change from 27.3 vs.27 ST and 24.8 vs.25.1 ST in the production of van Borsel et al.'s two subjects, and from 15.2 to 13.7 ST in Damrose's, is well within the ranges of documented time-of-day fluctuations (Lehto, Laaksonen, Vilkman, & Alku, 2008; van Mersbergen, Verdolini, & Titze, 1999), longitudinal fluctuations over a month (Gelfer, 1989), and on par with within-day vs.across-day differences (Murry, Brown, & Morris, 1995). As such, it indicates no perceptual narrowing of the available pitch range. On the contrary: the consistency of the size of the ranges is quite remarkable considering the time span of over one year between the first and last measurements. This reanalysis of the findings of previous studies indicates only a relocation of the entire pitch range to a lower fundamental frequency interval.

In addition, this dissertation documents the actual pitch ceiling movements during the transitioning of six FTM individuals. While the pitch floor for all subjects in the present study steadily lowered during the first year, the pitch ceiling indeed went through a lowering phase (even when calculated in ST). This phase happened approximately 6-7 months after starting testosterone. It manifested either in the loss of access to the original pitch ceiling, as in the case of CM, LW and NL (see Figure 3.10 above) or a plateau in the increasing pitch ceiling, such as in the case of MS and TJ. As evidenced by the same figure, by the end of the first year into the androgen therapy, all subjects fully regained or even surpassed their original pitch ceiling.

Based on the irreversible quality of testosterone, and in general androgen-induced, changes (e.g., Pattie et al., 1998), transmen were expected to show only downward movement of the habitual pitch. Upward movement was considered impossible on physiological bases.

The data from the six transmen did not support this hypothesis as three of them showed a late rise in their mean habitual pitch (referred to as a reverse J-pattern above). As the experiences of transwomen and women on Danazol so succinctly show, the lowering effects of testosterone on the larynx are irreversible. This means that the observed upward shift is not due to physiological reasons, but more likely individual choices transmen make. It seems that some FTMs choose to speak not only markedly higher than their predicted optimal pitch, but consciously or unconsciously choose to systematically shift their pitch within their accessible range.

The evidence that transmen may make choices that counter the physiologically predicted and originally highly desired effects of testosterone sheds light not only on the continuum between and around genders, but also the conscious or unconscious fine-tuning of vocal behavior transmen employ to display and perform their current sex and gender. This sort of fluidity, flux and fine-tuning has to be given space both in clinical practice and in the theoretical system of gender. Assuming that all transmen wish to perform unambiguously masculine traits incorrectly envisions transitioning as a binary switch and perpetuates a binary gender system as the only available gender options.

Chapter 4

The effects of androgen therapy on the filter function of the biologically female vocal tract

4.1 Introduction

4.1.1 Abstract

The previous chapter documents the changes in the voicing source in six transmen. After considering the voicing source, this chapter now turns to examining the testosterone-induced changes in the vocal tract filter function.

The results of this longitudinal study reveal previously undocumented changes in vowel formant values. The first three formant values, which were previously not considered a factor during the vocal transitioning of transmen, significantly lower during the first year of hormone therapy. The formant data results from two subjects followed for multiple years display similarities with the pattern of changes of the vocal source, inasmuch as the physiological changes can be partially counteracted by behavioural choices.

The changes documented in this chapter broaden the understanding of the effects of (cross-)sex hormones on vocal tract length. Firstly, the data from the current study challenge the theory that the longer vocal tract length (and as a result, lower formant values) in men compared to women is due to the male growth spurt in adolescence

and the general growth differential between the sexes. As transmen do not grow in height as a result of testosterone therapy, the lowering formant values may indicate either a lower(ed) laryngeal setting, general maxillary and mandibular growth that is possible well into the 20s and 30s of these individuals, or behavioural changes.

Secondly, the increased ability to make behavioural choices, which is facilitated by the physiologically altered structures, provides further evidence of the relative flexibility of speech features that were previously thought of as largely fixed. Slow fluctuation and sometimes reversal of the values also support the idea that speakers continually fine-tune features as they repeatedly re-locate themselves in the gender continuum.

4.1.2 Expectations based on previous literature

In previous studies on the effects of exogeneous testosterone on the female larynx, be it in cisgender or transgender individuals, the sole focus has always been on the vocal source with accounts invariably reporting lowering of the fundamental frequency. Discussions of the potential or actual effects of androgens on the vocal tract length, or explicitly on the formant values, is a topic curiously absent from the literature.

General formant issues in transgender speech therapy were considered sporadically in the 1980s and 1990s (R. Coleman & Markham, 1991; Günzburger, 1989; Mount & Salmon, 1988; Oates & Dacakis, 1997) but systematically only since around 2006 (Carew, Dacakis, & Oates, 2007; Davies & Goldberg, 2007), and in every case solely in the context of male-to-female individuals. The MTF bias often manifests by the authors referring to the entire transgender and transsexual population but discussing issues from an MTF vantage point, such as: "Previous research also suggests that formant frequency values in transsexuals may be changed to approximate those of female speech by altering oral resonance via increasing forward tongue carriage and lip spreading" (Carew et al., 2007, p. 593).

To exemplify the still prevailing scientific discourse, the chapter "Resonance" from the clinical guide *Voice and communication therapy for the transgender / transsexual client* (Adler & van Borsel, 2006) is summarized below.

Hirsch (2006, p. 209–210) starts with pointing out that terms such as "resonance" and "focus" are staple concepts in the musicology and clinical literature and quotes Peterson et al. (1994) as defining resonance as the basis of her approach in training MTFs: "Resonant voice is associated with the proprioception of oral vibratory sensations on or near the alveolar ridge and other facial plates, and in breathy voices, the auditory impression of ringing." In the article, Hirsch uses the terms "resonance", "resonant voice" and "tone focus" interchangeably.

Hirsch continues to suggest that "[Besides habitual pitch and intonation] Resonance may be considered the third portion of the vocal triumvirate that contributes toward a male / female *sound* differentiation." (2006, p. 210). Admitting that resonant quality (or tone) is influenced by formant values, she maintains resonance cannot be approached instrumentally, as "the signal-to-noise ratio on an average clinic recording system makes true measurement of this kind questionable at best" and "it has to be judged subjectively." The overlapping therapeutic approaches of "forward focus" and the "/i/-ification of vowels", while never explicitly stated in the chapter, clearly promote formant raising strategies such as lip-spreading and shortening of the vocal tract by way of the tongue pulling the larynx upward. However, Hirsch evaluates the voices resulting from said therapeutic techniques as applied to MTF patients along terms such as "masculine / chesty" and "feminine / upper larynx / head."

This succinctly illustrates the still-prevailing conflation of the uses of the word "resonance" in the literature. In the above example the acoustic definition of formant frequencies, i.e.,the filter function effects of the vocal tract on the source signal was inserted into and evaluated in a musicological discourse on the voice quality that corresponds to vibrations of the source signal in various cavernous parts of the body (as in head vs.middle vs.chest voice).

However, this practice introduces problems on multiple levels. In both therapy and research on transgender speech it makes formants appear to be a factor contributing to a phenomenon, resonance, that cannot be measured instrumentally. In therapy, the claim that resonance needs to be subjectively evaluated very likely discourages transgender patients from self-measuring and monitoring formant changes as milestones during their transitioning.*

In clinical research there has been little instrumental analysis carried out on TG formants (Carew et al. (2007), ten MTFs; Gelfer and Schofield (2000), 15 MTFs; Gunzburger (1989, 1993, 1995), six MTFs; Mount and Salmon (1988), one MTF). No work has been published on FTM formants yet.

---

*While working with members of the Transgender Foundation of America, it was my experience that after 2-3 sessions of group practice of working with Praat, both FTMs and MTFs expressed how they liked being able to get objective voice measurements of themselves.

4.2 Hypotheses to be tested

The aim of the present study was to investigate the effects of testosterone therapy on FTM formant values in a sample of seven female-to-male individuals.

Hypothesis 1. *Based on the expectation that transmen will not experience laryngeal lowering or a growth spurt, each vowel phoneme under consideration will retain its pre-testosterone values for F1, F2, and F3. That is, there will be no significant interaction of vowel with time on testosterone.*

4.3 Subjects

The speech of seven non-smoking, musically untrained female-to-male transgender individuals (mean age 32.3, min = 23, max = 47, c.f.Table 4.1 on page 74) undergoing androgen therapy was documented before the commencement of the testosterone therapy and one year into their hormonal transitioning from female to male.

In weekly instalments the speakers self-administered 75-250+ mg testosterone cypionate intramuscularly. However, subjects did not always use exact amounts and their dosage was often re-titrated according to their physiological and emotional needs. Therefore the cited dosage should be viewed as indicating the intended dosage and not as factual information on the subjects' actual testosterone level.

4.4 Materials and procedures

The reading material from the pitch tasks was also used in this study. The stressed monophthongs in the *Rainbow Passage*, the *Houston Urban English Survey Passage*,

Subject ID	age	available data (years)
CM	23	3
EX	31	5
JS	38	1
LW	38	2
MS	23	1
SB	30	1
TJ	32	1
average	30.7	

Table 4.1 : Subjects in the longitudinal study on formants

and *Comma Gets a Cure* were used to collect vowel space information. To enhance the representativeness of the vowel formant frequencies obtained, it was ensured that the words from which the vowels were extracted were not flanked by another vowel, liquid or rhotic as this would likely influence the formant values. Vowels in these environments were excluded from the analysis. As the same reading passages were used at every time point and the vowels were analysed paired by word, any coarticulatory influences would be expected to be similar for all measures across time.

During his self-documentation EX recorded the following four sentences at every time point: *"This is Ethan. It's [month] [day], [year]. I've been on T for [number of months / days]. I'll get my [nth] shot of testosterone today."* In this sample the high front corner vowel was represented at least twice in every session in the words *Ethan* and *T* and often in the numbers as well, such as *14th*, *2003*, etc. While the number of vowels does not allow us to form a complete picture of EX's changes, the temporal resolution of his longitudinal data (five years' worth of self-documentation) makes his contribution valuable.

Lexical set	Phonetic value	Examples
FLEECE	/iː/	*TV, jeep, need, repeated*
KIT	/ɪ/	*chip, stick, big, pick, figure*
FACE	/eɪ/	*day, game, Jays, date, maybe*
DRESS	/ɛ/	*said, dead, ever, event*
TRAP	/æ/	*match, happens, matter, actions*
LOT	/ɔ or ɑ or ɒ/	*bother, box, hockey, thought, Hawks, Dawn, talk*
STRUT	/ʌ/	*up, puck, shrugged, something*
FOOT	/ʊ/	*put, book, cookies, hood*
GOOSE	/uː/	*food, dude, boots, Sue, juice*

Table 4.2 : Lexical sets and examples for the vowel categories used in the formant analysis

4.5 Analysis

Formants were measured in a representative sample of vowels from the reading passages. The acoustic analysis was carried out on the data using Praat ver. 5.1.02 and a manual formant logging script package written by Phil Harrison. The nine vowel categories (listed in Table 4.2 on page 75) follow the lexical sets established by Wells (1982).†

Formant measurements for F1, F2 and F3 were averaged over relatively stable

†Wells (1982, p. 123) explains his choice of keywords (KIT, FLEECE, FACE, etc.) as follows:

> "the keywords have been chosen in such a way that clarity is maximized: whatever accent of English they are spoken in, they can hardly be mistaken for other words. Although *fleece* is not the commonest of words, it cannot be mistaken for a word with some other vowel; whereas *beat*, say, if we had chosen it instead, would have been subject to the drawback that one man's pronunciation of beat may sound like another's pronunciation of *bait* or *bit*."

Subject	Year	Max formant value (Hz)	No. of formants	No. of vowel tokens
CM	0	5000	4.5	215
CM	1	4600	4-4.5	286
CM	2	4000	4	277
CM	3	4500	4	281
JS	0	4000	4	280
JS	1	4000	4	359
LW	0	5000	4.5	287
LW	1	5000	4.5	290
LW	2	5000	4.5-5	275
MS	0	5500	4.5	285
MS	1	4300	4.5	278
SB	0	4500	4-4.5	239
SB	1	4200	4	239
TJ	0	4500	4.5	278
TJ	1	4500	4	281

Table 4.3 : Formant tracker settings in Praat by recording session. 0 indicates the session before the commencement of the testosterone therapy. 1 marks the recording session one year after the commencement, etc.

intervals at the steady state of each vowel. Voiceless, overshort, creaky tokens, as well as tokens with untrackable first three formants were excluded from the analysis. Between 10 and 50 tokens were analysed per vowel per speaker per time point, yielding a total of 250-300 measured tokens per recording session. The Burg method of analysis used a window length of 0.025 s, 30 dB dynamic range and 50 Hz pre-emphasis. The rest of the formant analysis settings were re-calibrated session to session to accurately track formants. For the settings in each session see Table 4.3 (on page 76).

For subject EX, every available token of the high front vowel /iː/ was measured. In the first year this was 127 tokens, the second 26, the third 5, and the fifth 3. His data are not interpreted in yearly increments but as a general trend calculated from his session-to-session data.

In the production of many subjects, the vowel in the FACE lexical category was largely monophthongal even in reading passages. For subjects who did monophthongize this vowel, it was treated and measured as a monophthong and these subjects had nine vowel categories measured. For subjects with a diphthongal FACE vowel, eight vowel categories (/iː, ɪ, ɛ, æ, ɑ, ʌ, ʊ, and uː/) were measured and plotted.

Subjects also differed with regards to their production of the so-called low back merger, that is, the merger of the vowels in word pairs such as DON-DAWN and COT-CAUGHT. There is evidence in dialect surveys (Niedzielski & Koops, 2011; Labov, Ash, & Boberg, 2005) that Texas in general shows both unmerged and variably merged tokens of these lexical sets, so in this study the three canonical low back vowels, /ɔ/, /ɑ/ and /ɒ/, were treated as one vowel category, /ɑ/.

4.6 Results

4.6.1 Change in formant values F1-3 in the first year on testosterone

Three mixed model analyses of variance (ANOVA) were conducted with the within-subjects repeated measure being time on testosterone (two levels indicating the number of years, zero or one) and the between-subjects factor being vowel quality. The dependent variables were each of the measured F1-3 values. Vowel token measurements from multiple sessions of the same speaker were paired by word and vowel (cf.Table 4.4 on page 78).

The mixed model ANOVAs indicated a significant interaction effect between testos-

F1(0)	F2(0)	F3(0)	Vowel	Word	F1(1)	F2(1)	F3(1)
576	1890	3253	/ɛ/	Ben	516	1831	2962
455	1528	2750	/ʊ/	put	438	1593	2727
446	1514	2734	/ʊ/	book	465	1392	2555
800	1277	2559	/ɑ/	thought	700	1286	2490
379	2510	2942	/iː/	he	373	2540	3051
420	1854	3051	/ɪ/	sitting	401	1823	2900
...	...	...	...	...	...	...	...

Table 4.4 : Examples of the paired data from JS's pre-testosterone session (F1(0), F2(0), F3(0)) and the session one year into the hormone therapy (F1(1), F2(1), F3(1)) in the statistical software package.

terone and vowel quality in all subjects but usually not for all formants (for the complete ANOVA results see Table 4.5 on page 83). The formant lowering effect can be observed in virtually every vowel in F1 in every subject, however, open vowels seem to be affected by it more than mid or close vowels. For figures illustrating the interaction effects see Figures 4.1-4.3. The lowering shift in F2 was fairly evenly distributed regardless of vowel quality in all six subjects. Subjects LW, MS, JS and SB exhibited lowering of F3 in virtually every vowel. For LW, the pronounced lowering affected every vowel in a uniform manner. In the case of MS, JS and SB, while there was a lowering effect on virtually every vowel in their repertoire, close vowels were affected more than mid or open vowels. CM and TJ exhibited changes with no easily describable pattern. For CM, in vowels /eɪ, ɛ, æ/ F3 formant values were *higher* after one year on testosterone, and in the other vowels they were lower. For TJ, vowels /ɪ, æ/ show virtually no change. Vowels /ʊ, uː/ showed *higher* F3 values after one year on testosterone, while the F3 in other vowels lowered according to expectations.

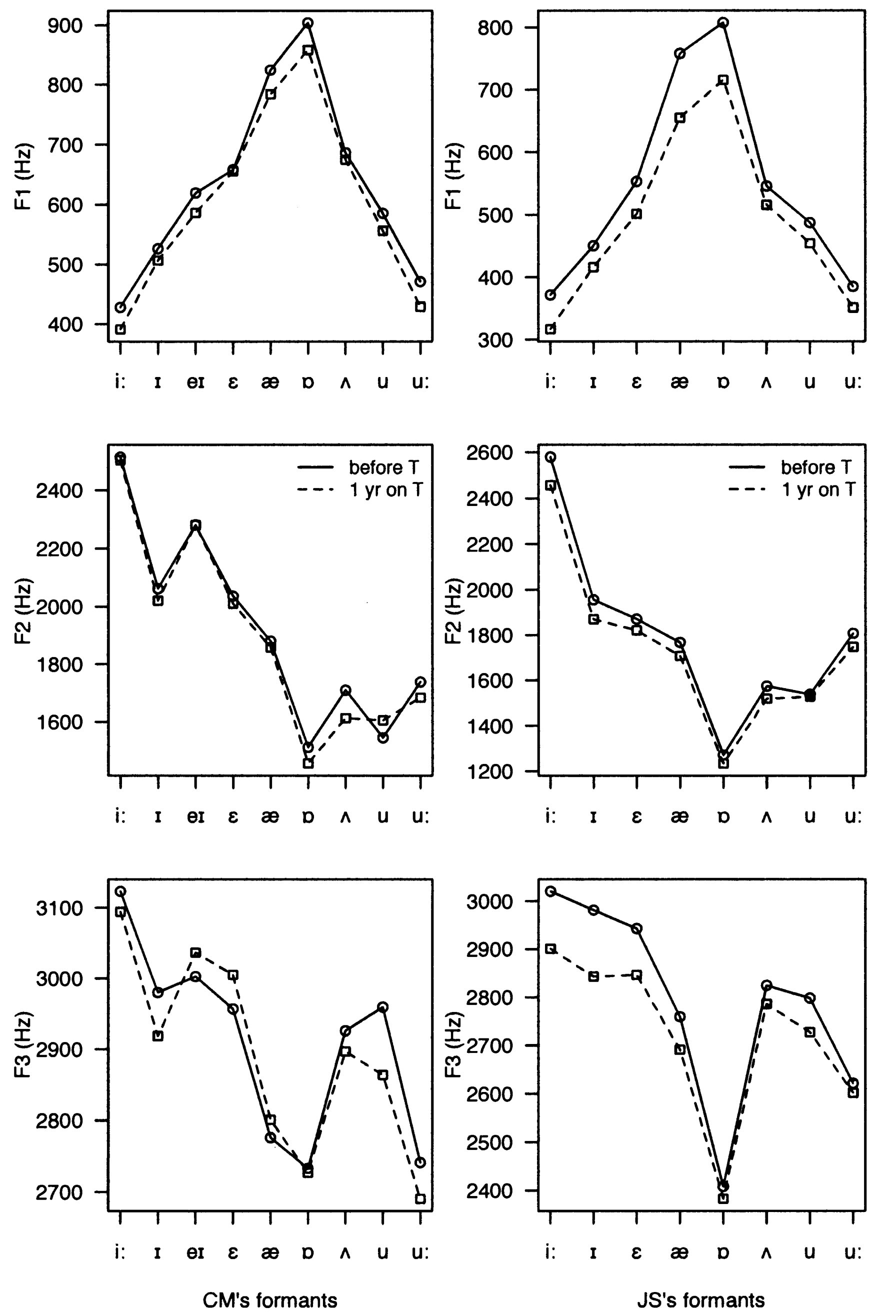

Figure 4.1 : Interaction between testosterone and vowel quality over 12 months for subjects CM and JS.

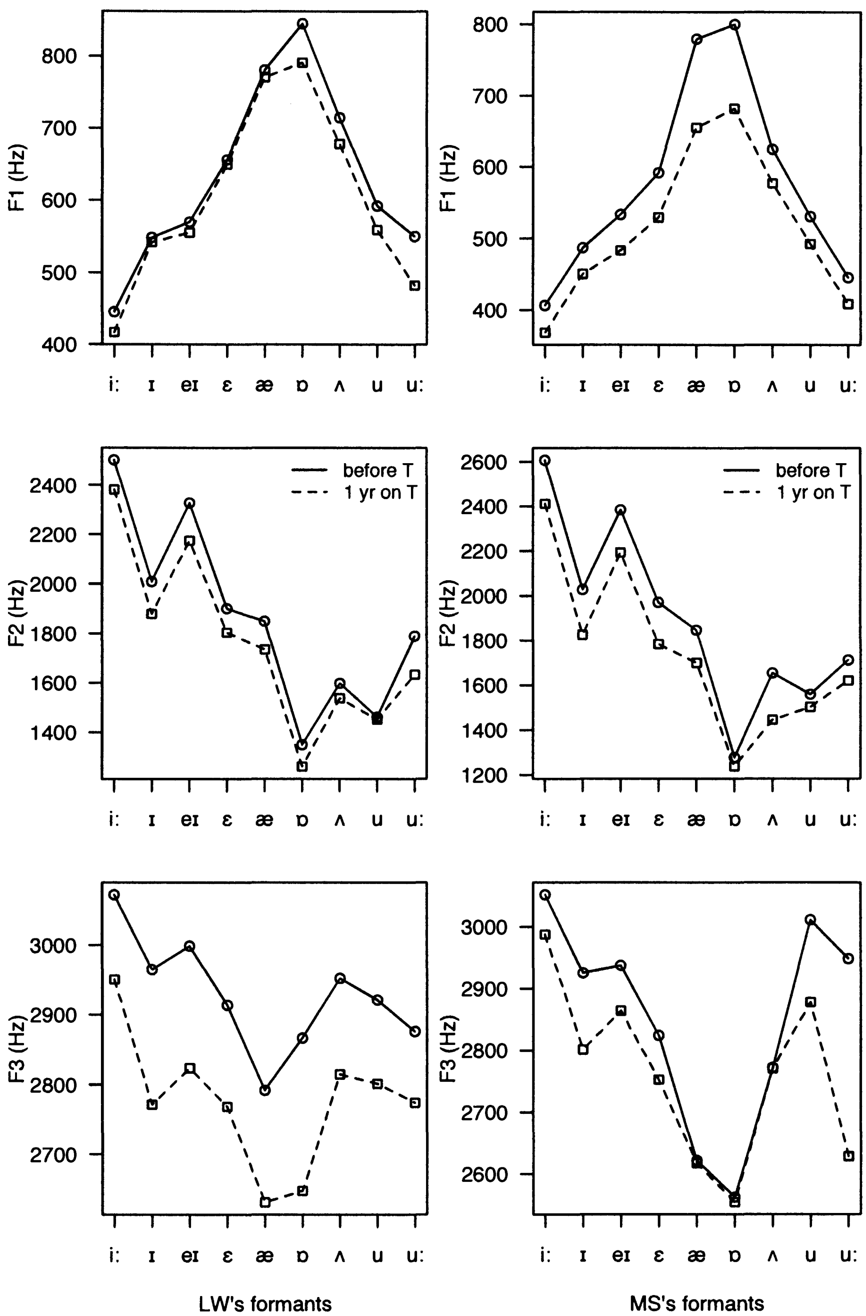

Figure 4.2 : Interaction between testosterone and vowel quality over 12 months for subjects LW and MS.

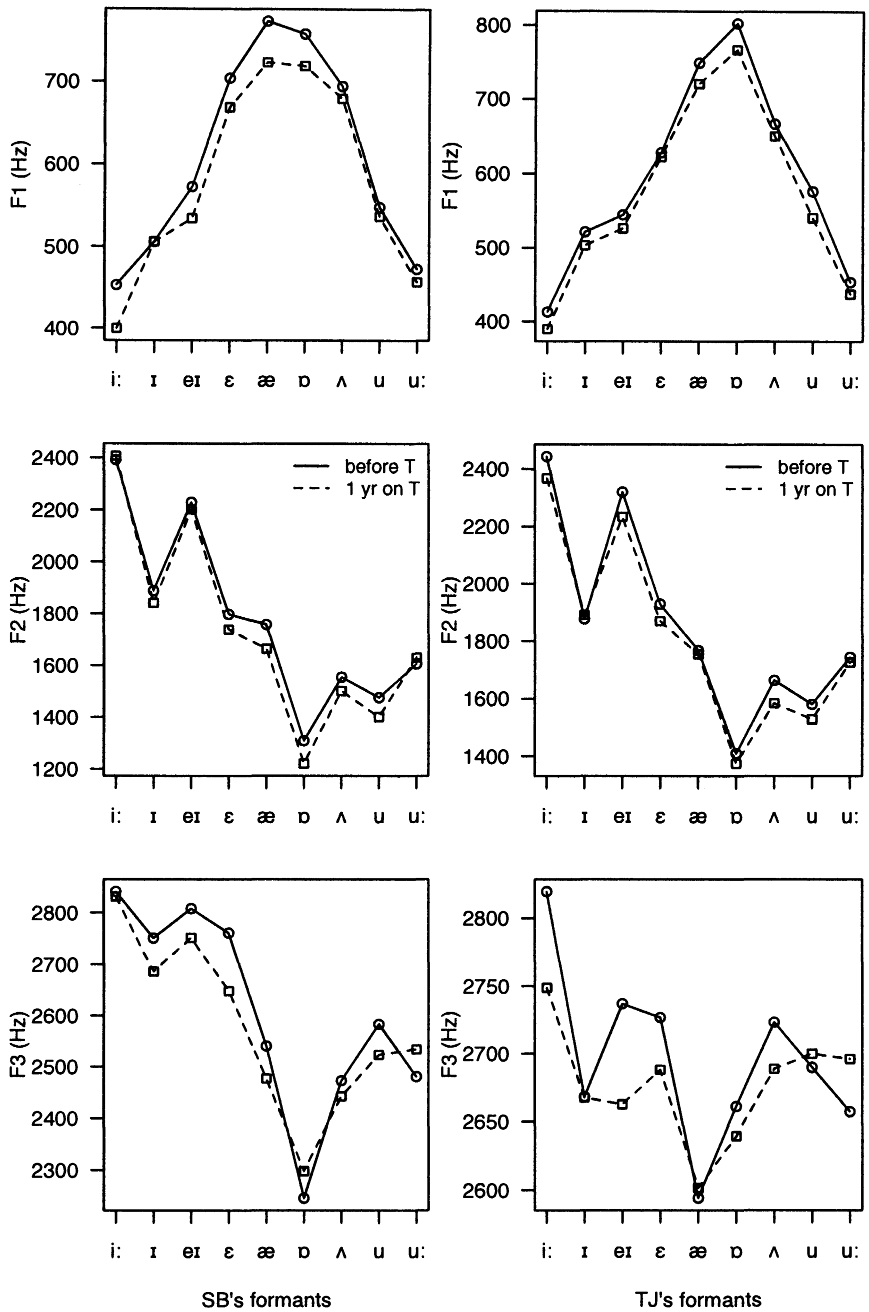

Figure 4.3 : Interaction between testosterone and vowel quality over 12 months for subjects SB and TJ.

The mixed model ANOVAs also indicated a significant main effect of testosterone on F1 and F2 for all six subjects at the alpha level of 0.05 (for the complete ANOVA results see Table 4.5). A significant main effect of testosterone was found on F3 in four subjects, with CM and SB as exceptions. Transmen exhibited lowered formant values by the end of their first year on testosterone.

Vowel also showed a significant main effect in all formants in all speakers. This phenomenon was expected, however, since the vowel qualities differ precisely in formant values. For scatterplots illustrating the changes over one year on the traditional F1 $\times$ F2 plane see Figures 4.4-4.9.

Clearly, these results establish that contrary to expectations, the filter function of transmen's vocal tract changed over time. The direction and magnitude of the change in the formant values indicate that the vocal tract got longer even though the transmen themselves did not grow taller. As laryngeal lowering was not measured in these subjects, its effect on the formant values in this study cannot be ruled out, nor can the effects of behavioural changes to the articulators.

While the original goal of this study was to document the formant values until the end of the first year on testosterone, after the changes became apparent, questions arose about the time course of the changes as well as whether they are unidirectional. Three subjects in the study agreed to provide recordings past the one year mark. The next section uses these data to investigate both questions.

subject	formant	main effect of testosterone	main effect of vowel	interaction effect of T * V
CM	F1	F $(1, 258) = 53.18$, $p < 0.001$	F $(8, 258) = 182.23$, $p < 0.001$	F $(8, 258) = 2.61$, $p = 0.009$
	F2	F $(1, 258) = 22.86$, $p < 0.001$	F $(8, 258) = 110{,}63$, $p < 0.001$	*ns* (F $(8, 258) = 1.64$, $p = 0.112$)
	F3	*ns* (F $(1, 258) = 1.38$, $p = 0.242$)	F $(8, 258) = 24.21$, $p < 0.001$	*ns* (F $(8, 258) = 1.23$, $p = 0.281$)
JS	F1	F $(1, 254) = 190.67$, $p < 0.001$	F $(7, 254) = 289.55$, $p < 0.001$	F $(7, 254) = 11.72$, $p < 0.001$
	F2	F $(1, 254) = 50.11$, $p < 0.001$	F $(7, 254) = 376.02$, $p < 0.001$	F $(7, 254) = 3.07$, $p = 0.004$
	F3	F $(1, 254) = 28.48$, $p < 0.001$	F $(7, 254) = 49.01$, $p < 0.001$	F $(7, 254) = 2.17$, $p = 0.037$
LW	F1	F $(1, 255) = 25.58$, $p < 0.001$	F $(8, 255) = 122.54$, $p < 0.001$	*ns* (F $(8, 255) = 1.91$, $p = 0.59$)
	F2	F $(1, 255) = 186.2$, $p < 0.001$	F $(8, 255) = 115.98$, $p < 0.001$	F $(8, 255) = 3.09$, $p = 0.002$
	F3	F $(1, 255) = 235.22$, $p < 0.001$	F $(8, 255) = 20.7$, $p < 0.001$	F $(8, 255) = 2.89$, $p = 0.004$
MS	F1	F $(1, 232) = 166.94$, $p < 0.001$	F $(8, 232) = 136.91$, $p < 0.001$	F $(8, 232) = 10.15$, $p < 0.001$
	F2	F $(1, 232) = 211.28$, $p < 0.001$	F $(8, 232) = 115$, $p < 0.001$	F $(8, 232) = 6.89$, $p < 0.001$
	F3	F $(1, 232) = 22.97$, $p < 0.001$	F $(8, 232) = 26.68$, $p < 0.001$	F $(8, 232) = 2.42$, $p = 0.016$
SB	F1	F $(1, 189) = 77.17$, $p < 0.001$	F $(8, 189) = 127.75$, $p < 0.001$	F $(8, 189) = 4.06$, $p < 0.001$
	F2	F $(1, 189) = 27.48$, $p < 0.001$	F $(8, 189) = 92.41$, $p < 0.001$	F $(8, 189) = 3.55$, $p = 0.001$
	F3	*ns* (F $(1, 189) = 2.53$, $p = 0.114$)	F $(8, 189) = 26.73$, $p < 0.001$)	*ns* (F $(8, 189) = 1.27$, $p = 0.26$)
TJ	F1	F $(1, 246) = 42.13$, $p < 0.001$	F $(8, 246) = 194.2$, $p < 0.001$	*ns* (F $(8, 246) = 1.46$, $p = 0.17$)
	F2	F $(1, 246) = 32.84$, $p < 0.001$	F $(8, 246) = 124.31$, $p < 0.001$	*ns* (F $(8, 246) = 1.97$, $p = 0.51$)
	F3	F $(1, 246) = 5.63$, $p < 0.001$	F $(8, 246) = 7.17$, $p < 0.001$	F $(8, 246) = 2.28$, $p = 0.023$

Table 4.5 : ANOVA results for the first three formants.

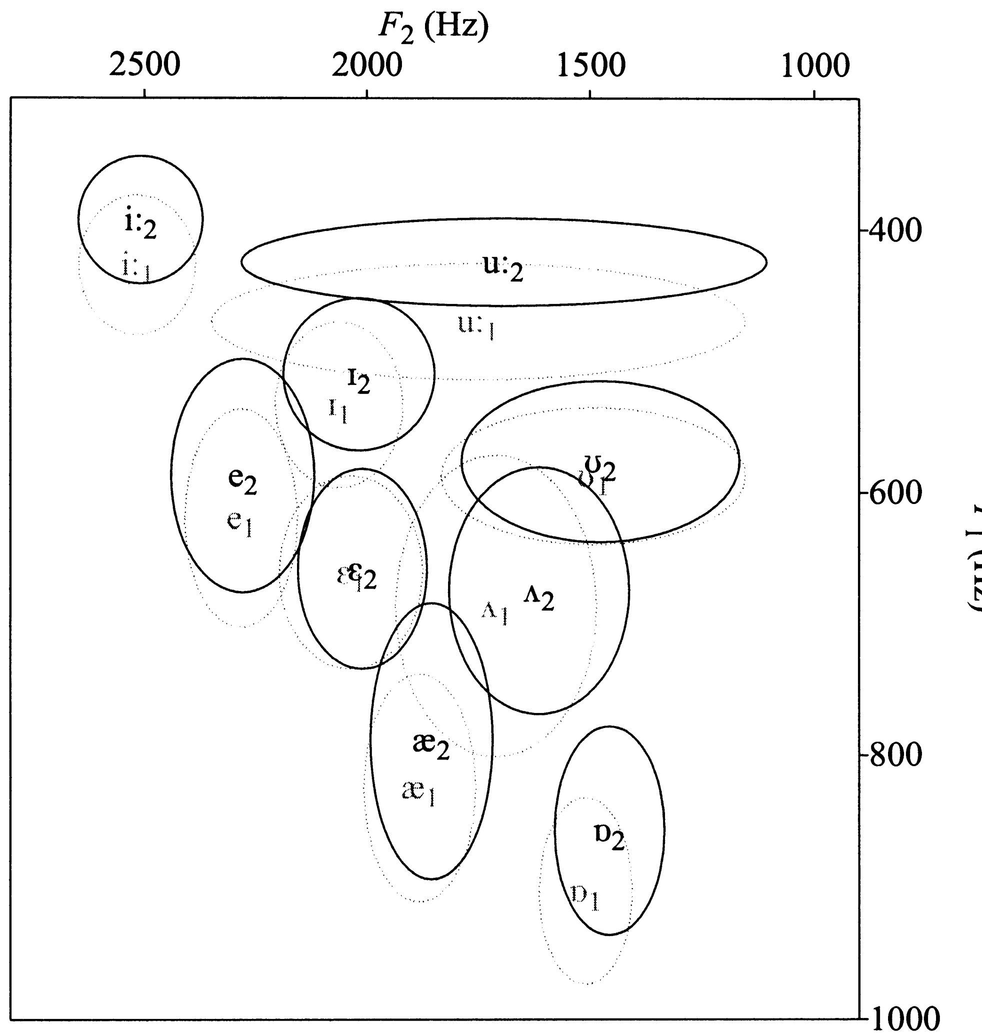

Figure 4.4 : F1 × F2 plot of CM over 12 months. Grey: pre-testosterone values; Black: values after 12 months of testosterone therapy; Ellipses: one standard deviation from the mean.

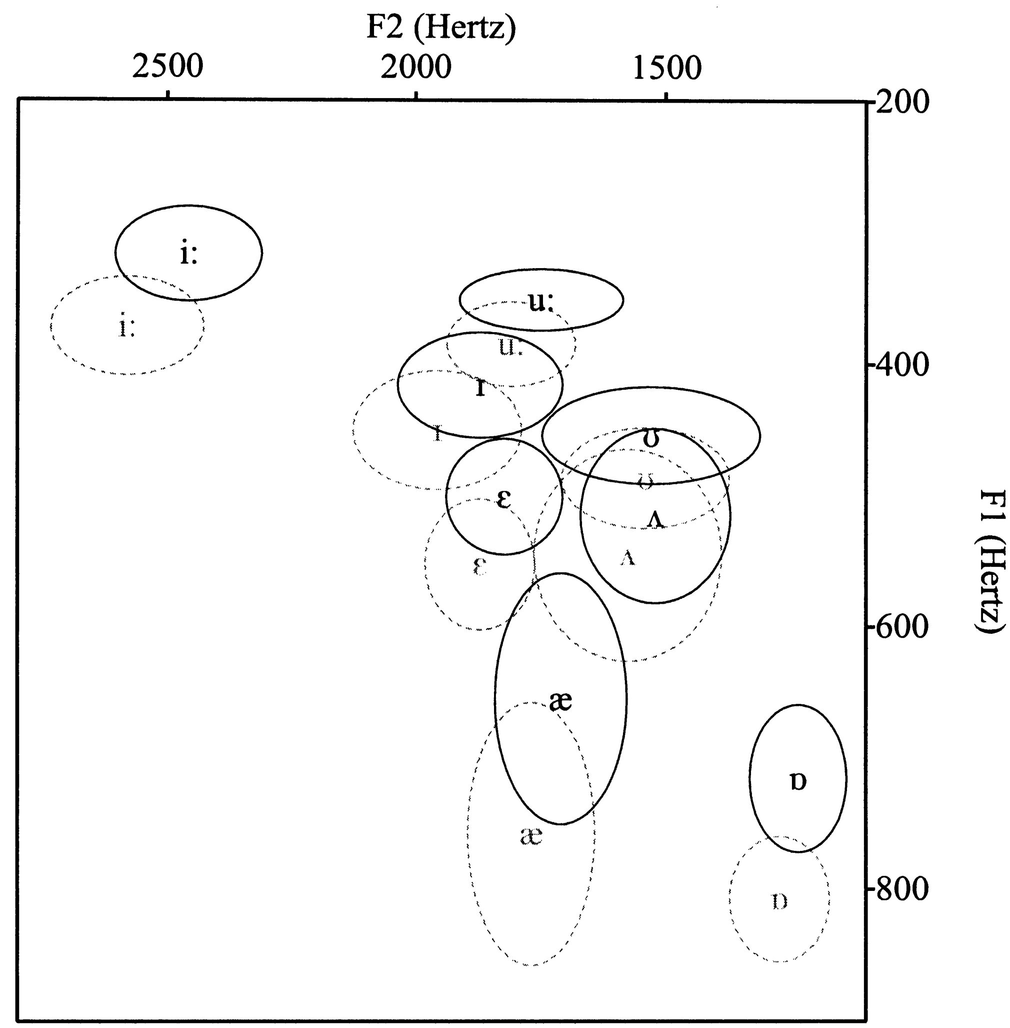

Figure 4.5 : F1 × F2 plot of JS over 12 months. Grey: pre-testosterone values; Black: values after 12 months of testosterone therapy; Ellipses: one standard deviation from the mean.

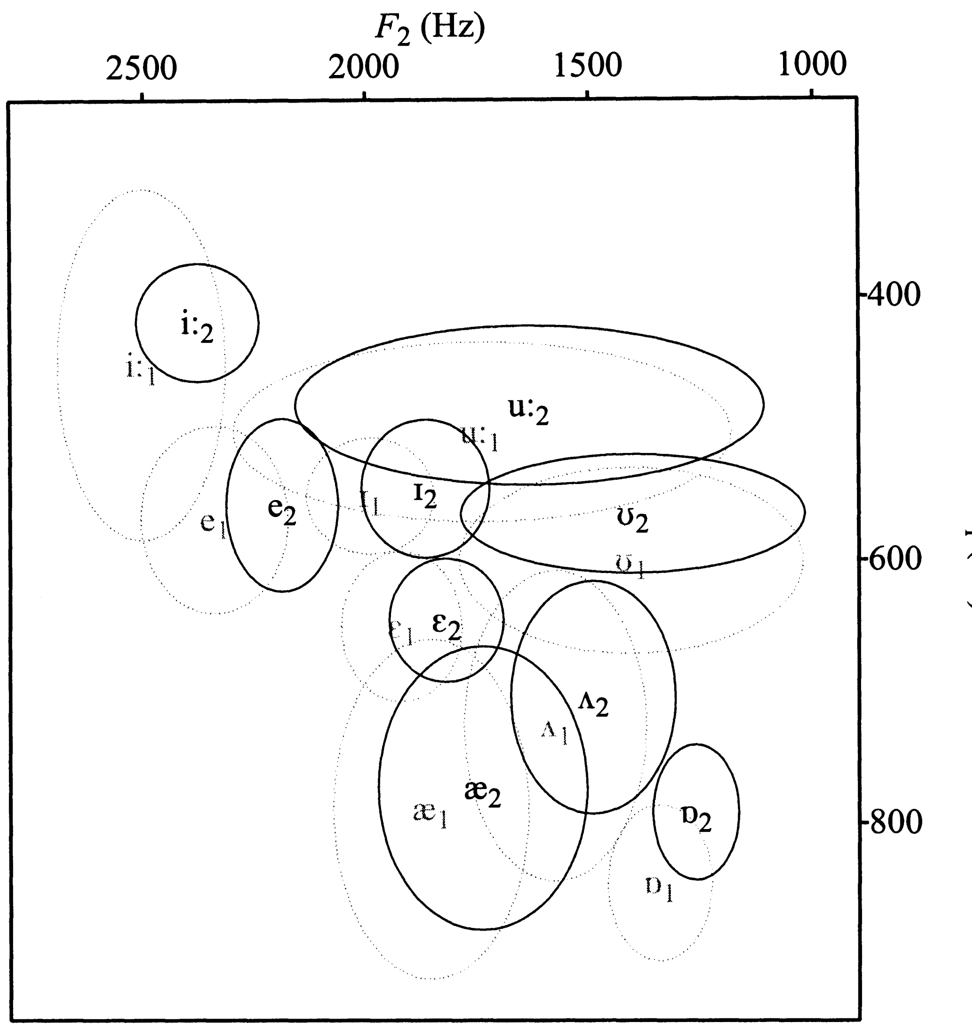

Figure 4.6 : F1 × F2 plot of LW over 12 months. Grey: pre-testosterone values; Black: values after 12 months of testosterone therapy; Ellipses: one standard deviation from the mean.

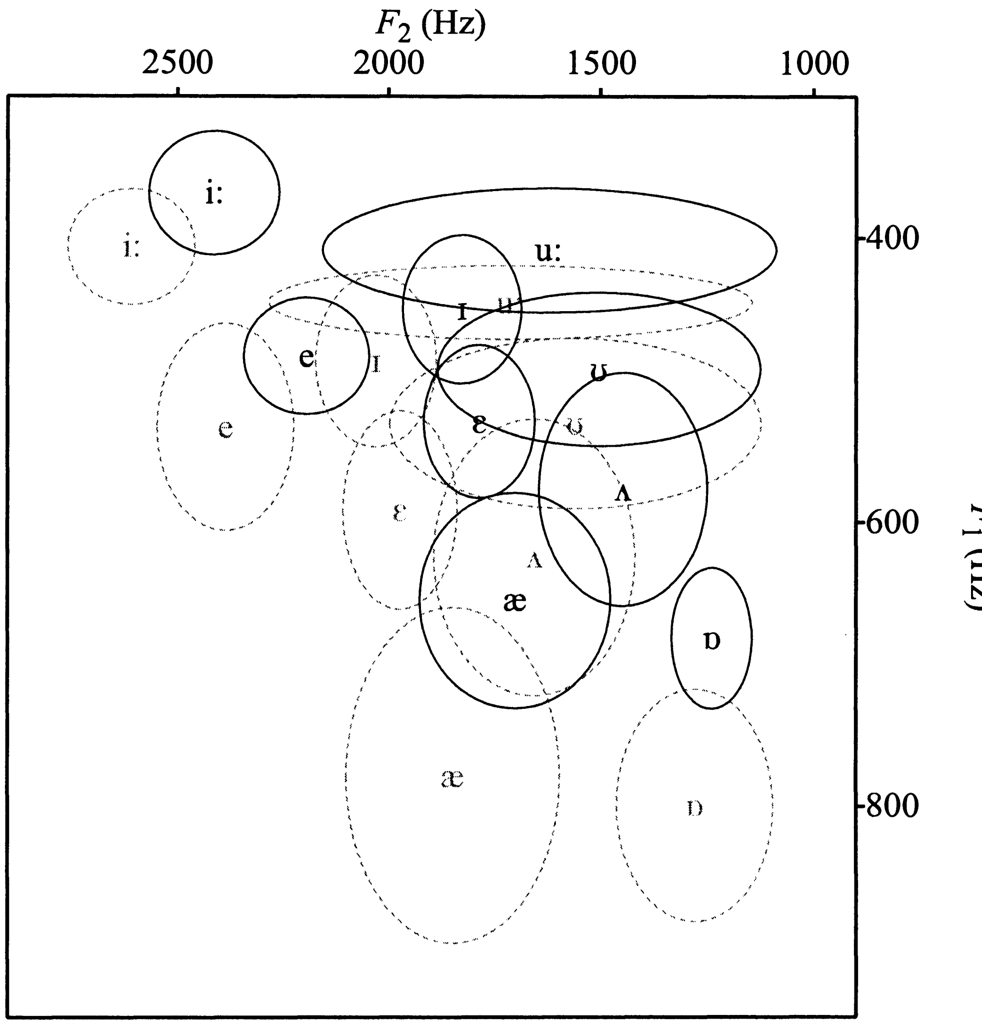

Figure 4.7 : F1 × F2 plot of MS over 12 months. Grey: pre-testosterone values; Black: values after 12 months of testosterone therapy; Ellipses: one standard deviation from the mean.

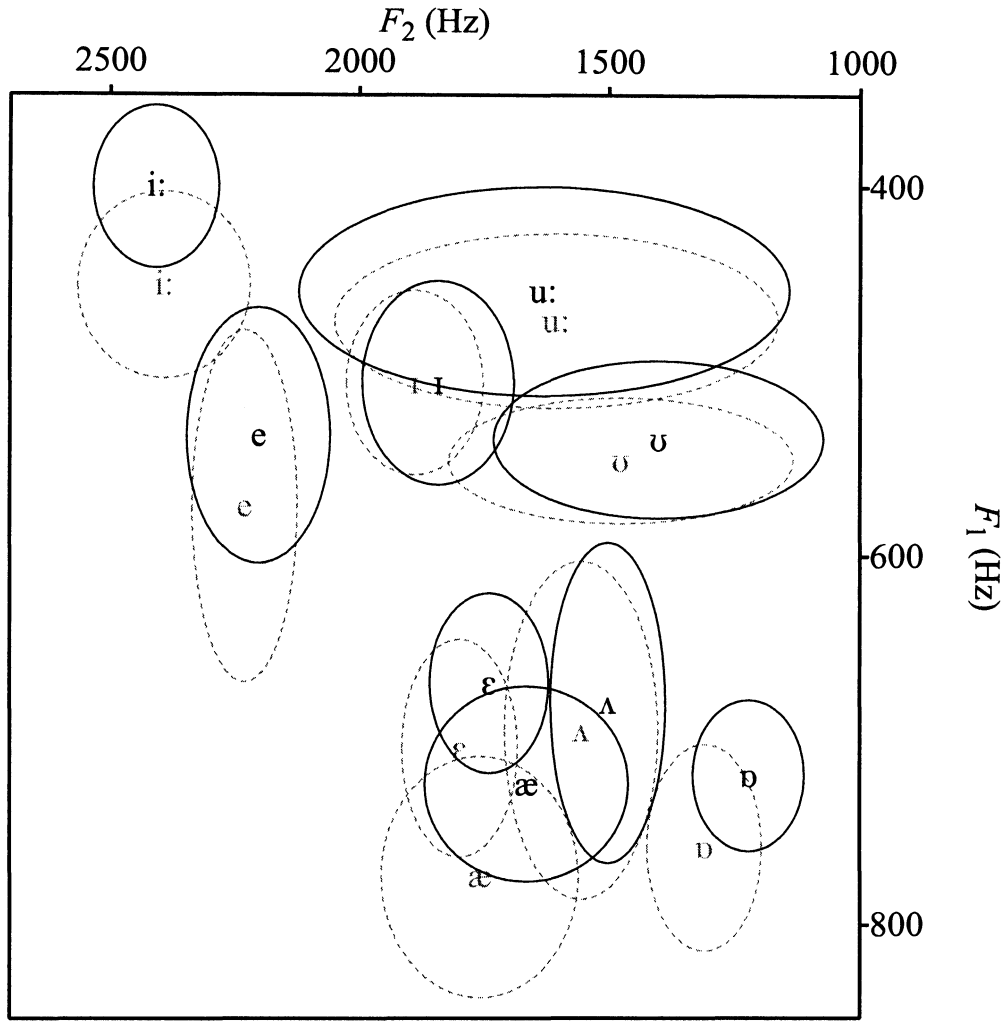

Figure 4.8 : F1 × F2 plot of SB over 12 months. Grey: pre-testosterone values; Black: values after 12 months of testosterone therapy; Ellipses: one standard deviation from the mean.

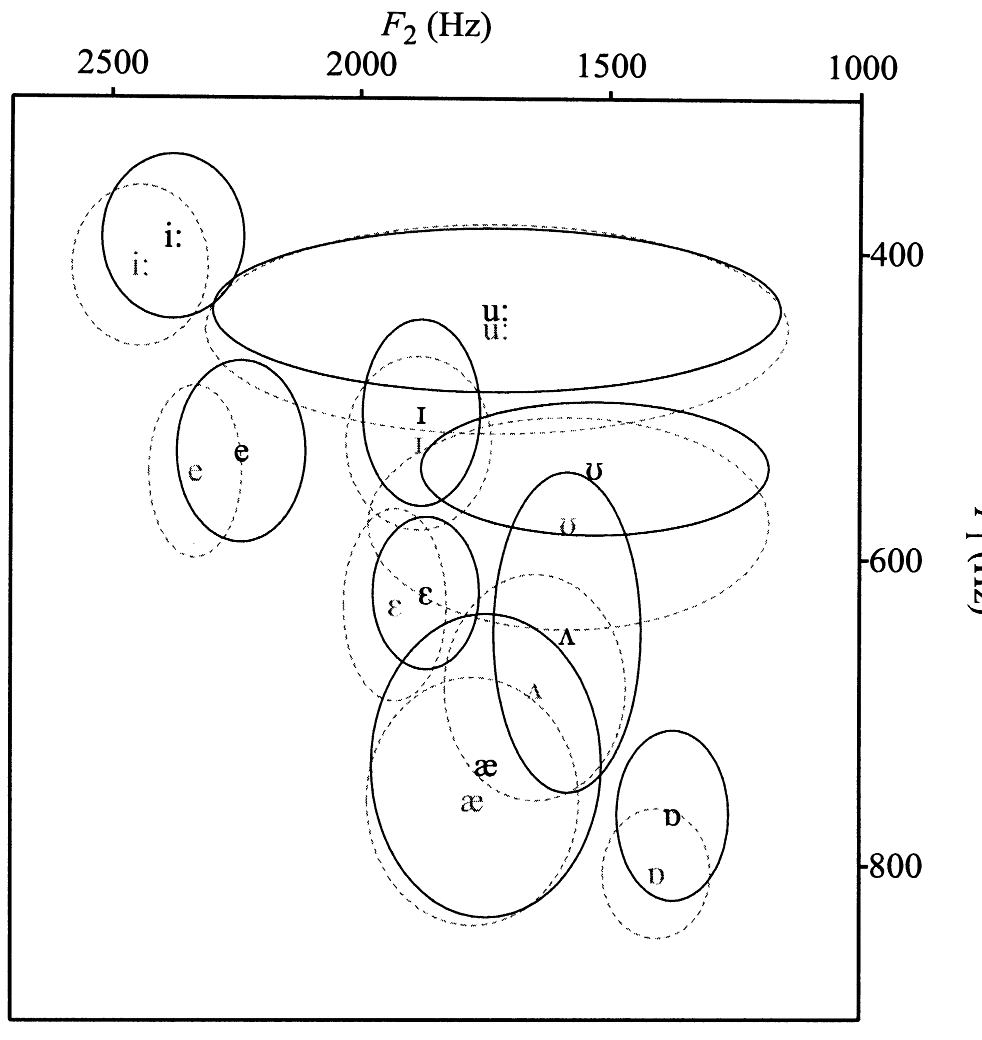

Figure 4.9 : F1 × F2 plot of TJ over 12 months. Grey: pre-testosterone values; Black: values after 12 months of testosterone therapy; Ellipses: one standard deviation from the mean.

4.6.2 Changes past the first year on testosterone

Two subjects, CM and LW, agreed to continue being documented past the first year mark. As of early 2011, three years of data are available for CM, and two for LW. As was previously mentioned, subject EX self-documented for approximately 5 years during his testosterone therapy. The data of these three subjects were used as a post-hoc study looking at long-term effects of testosterone on formant values.

As in the one year dataset above, two-way mixed-model repeated measures analyses of variance (ANOVA) were conducted with the within-subjects factor being time on testosterone (four levels for CM and three for LW, indicating the number of years on testosterone, zero to three), the between-subjects factor again being vowel quality. The dependent variable were the measured F1-3 values. As the significant main effect of vowel was predictable it is not reported in this section.

A first investigation of the data revealed different means for the formant measures for every yearly measurement. The means and standard deviations of CM and LW are summarized in Tables 4.7 and 4.6, respectively. In the charts "0" is used to index pre-testosterone values.

The ANOVA results of within-subjects effects of time on testosterone and the interaction effect between vowel quality and time on testosterone are presented in Table 4.8 (on page 93).

From Table 4.8 it follows that, apart from LW's F1, all the changes over the years on testosterone are statistically significant. When the year-to-year changes are investigated, it is also apparent that they do not stop in the first year and are

Vowel	N	F1(0)	SD	F1(1)	SD	F1(2)	SD
/æ/	35	790	(129)	774	(108)	776	(108)
/ɑ/	54	846	(59)	792	(51)	816	(60)
/eɪ/	15	569	(71)	558	(66)	564	(72)
/ɛ/	31	651	(57)	646	(47)	640	(54)
/ɪ/	18	551	(44)	546	(53)	542	(43)
/iː/	47	451	(134)	419	(45)	436	(156)
/ʊ/	13	601	(71)	565	(45)	588	(70)
/uː/	14	503	(68)	482	(60)	598	(430)
/ʌ/	18	726	(118)	705	(88)	715	(82)

Vowel	N	F2(0)	SD	F2(1)	SD	F2(2)	SD
/æ/	35	1855	(219)	1739	(234)	1763	(223)
/ɑ/	54	1344	(116)	1263	(96)	1255	(103)
/eɪ/	15	2340	(165)	2189	(126)	2200	(144)
/ɛ/	31	1920	(136)	1820	(128)	1840	(141)
/ɪ/	18	1990	(144)	1866	(143)	1882	(144)
/iː/	47	2504	(189)	2378	(138)	2409	(151)
/ʊ/	13	1407	(383)	1403	(384)	1381	(386)
/uː/	14	1741	(558)	1634	(525)	1692	(635)
/ʌ/	18	1576	(203)	1492	(183)	1484	(181)

Vowel	N	F3(0)	SD	F3(1)	SD	F3(2)	SD
/æ/	35	2782	(160)	2623	(152)	2639	(161)
/ɑ/	54	2857	(168)	2643	(156)	2782	(159)
/eɪ/	15	3005	(94)	2834	(64)	2830	(86)
/ɛ/	31	2927	(97)	2781	(108)	2777	(104)
/ɪ/	18	2961	(78)	2771	(95)	2794	(87)
/iː/	47	3077	(141)	2947	(131)	2987	(140)
/ʊ/	13	2941	(157)	2811	(163)	2775	(167)
/uː/	14	2839	(198)	2774	(174)	2766	(329)
/ʌ/	18	2964	(135)	2821	(173)	2830	(135)

Table 4.6 : Means and standard deviations of F1-3 in LW's production over 2 years.

Vowel	N	**F1(0)**	SD	**F1(1)**	SD	**F1(2)**	SD	**F1(3)**	SD
/æ/	33	824	(86)	788	(105)	766	(97)	772	(102)
/ɑ/	55	903	(71)	857	(79)	806	(60)	807	(66)
/eɪ/	28	619	(83)	587	(89)	549	(72)	560	(65)
/ɛ/	36	661	(74)	658	(76)	622	(74)	611	(76)
/ɪ/	21	533	(63)	510	(58)	510	(58)	495	(52)
/iː/	38	426	(53)	392	(48)	391	(49)	381	(52)
/ʊ/	13	587	(52)	576	(61)	563	(41)	554	(67)
/uː/	12	470	(44)	424	(33)	437	(47)	421	(39)
/ʌ/	21	686	(115)	674	(94)	638	(82)	648	(85)

Vowel	N	**F2(0)**	SD	**F2(1)**	SD	**F2(2)**	SD	**F2(3)**	SD
/æ/	33	1880	(125)	1854	(137)	1762	(143)	1787	(145)
/ɑ/	55	1509	(104)	1456	(122)	1361	(111)	1400	(126)
/eɪ/	28	2282	(127)	2279	(161)	2297	(164)	2284	(136)
/ɛ/	36	2035	(161)	2009	(145)	1906	(148)	1945	(156)
/ɪ/	21	2062	(143)	2017	(170)	1930	(163)	1992	(136)
/iː/	38	2515	(131)	2508	(139)	2488	(120)	2534	(126)
/ʊ/	13	1492	(339)	1477	(310)	1447	(290)	1441	(312)
/uː/	12	1750	(597)	1693	(588)	1677	(559)	1697	(577)
/ʌ/	21	1710	(225)	1613	(201)	1537	(188)	1570	(182)

Vowel	N	**F3(0)**	SD	**F3(1)**	SD	**F3(2)**	SD	**F3(3)**	SD
/æ/	33	2776	(160)	2798	(182)	2701	(170)	2774	(180)
/ɑ/	55	2727	(260)	2726	(201)	2554	(156)	2723	(217)
/eɪ/	28	2997	(123)	3029	(93)	2999	(146)	2994	(150)
/ɛ/	36	2954	(95)	3003	(134)	2901	(173)	2917	(116)
/ɪ/	21	2977	(110)	2923	(124)	2888	(121)	2950	(134)
/iː/	38	3130	(219)	3102	(208)	3085	(200)	3120	(208)
/ʊ/	13	2965	(180)	2894	(236)	2890	(159)	3019	(171)
/uː/	12	2782	(246)	2727	(238)	2724	(216)	2886	(268)
/ʌ/	21	2926	(190)	2897	(176)	2806	(167)	2865	(160)

Table 4.7 : Means and standard deviations of F1-3 in CM's production over 3 years.

	formant	main effect of testosterone	interaction effect of T * V
CM	F1	F (3, 248) = 72.83, $p < 0.001$	F (24, 248) = 3.56, $p < 0.001$
	F2	F (3, 248) = 52.52, $p < 0.001$	F (24, 248) = 4.45, $p < 0.001$
	F3	F (3, 248) = 15.83, $p < 0.001$	F (24, 248) = 2.48, $p < 0.001$
LW	F1	F (2, 236) = 4.57, $p = 0.01$	*ns* (F (16, 236) = 1.35, $p = 0.16$)
	F2	F (2, 236) = 110.38, $p < 0.001$	F (16, 236) = 1.91, $p = 0.017$
	F3	F (2, 236) = 133.67, $p < 0.001$	F (16, 236) = 3.30, $p < 0.001$

Table 4.8 : Tests of within-subjects effects for CM and LW.

statistically significant in later years as well. For CM the bulk of the changes happened in the later years, while LW completed the majority of the changes in his first year on testosterone.

In order to investigate the temporal relationships in the variables, the adjacent time points were examined with Bonferroni-corrected pairwise comparisons. For the results, see Table 4.9 and Table 4.10 (on page 94).

In CM's production, nine such comparisons were made (three formants × three time steps) using Bonferroni adjusted alpha levels of 0.006 (0.05/9). Five of these comparisons were significant at the corrected alpha level. In F1, the changes during year 1 and 2, in F2 the changes in the second year, and in F3 the changes in year 2 and 3 were significant. In LW's case, six such comparisons were made using the adjusted level of 0.0083 (0.05/6). In his case, all changes in the first year were significant, and F3 was significant in the second year as well.

Closer examination reveals that in their respective last year of the vocal documentation (second for LW and third for CM) the changes that both subjects exhibited were going the *opposite* direction from previous years. This pattern is very regular:

subject	formant	pairwise comparison	significance
CM	F1	before vs.1 yr	$p < 0.001$
		1 yr vs.2 yrs	$p = 0.004$
		2 yrs vs.3 yrs	$p = 1$
	F2	before vs.1 yr	$p = 1$
		1 yr vs.2 yrs	$p < 0.001$
		2 yrs vs.3 yrs	$p = 0.186$
	F3	before vs.1 yr	$p = 1$
		1 yr vs.2 yrs	$p < 0.001$
		2 yrs vs.3 yrs	$p < 0.001$

Table 4.9 : Pairwise comparisons by year in CM's production; Bonferroni-corrected alpha level = 0.006 for a familywise error rate of 0.05.

subject	formant	pairwise comparison	significance
LW	F1	before vs.1 yr	$p = 0.001$
		1 yr vs.2 yrs	$p = 0.148$
	F2	before vs.1 yr	$p < 0.001$
		1 yr vs.2 yrs	$p = 1$
	F3	before vs.1 yr	$p < 0.001$
		1 yr vs.2 yrs	$p = 0.006$

Table 4.10 : Pairwise comparisons by year in LW's production; Bonferroni-corrected alpha level = 0.0083 for a familywise error rate of 0.05.

it applies to both speakers, for virtually every vowel. In order to illustrate this, the multi-year changes of the two subjects are plotted in Figures 4.10 and 4.11 .

4.7 Discussion

Although the participants did not get taller, the mean differences observed in the ANOVA models show clearly that formant values lowered significantly with exposure to testosterone and thereby the filter function of transmen virilizes. Over 12 months the means in all three formant values become lower with statistically significant change achieved by all subjects by the end of the first year. Evidence from two case studies suggests that sometime during the second or third year on testosterone the change in formant values may reverse, which results in values rising.

Although these results run counter to my hypothesis, the data available at this point do not allow us to ascertain whether the changes in formant values in FTMs are purely physiologically conditioned or whether behavioural and psycho-acoustic considerations also feed into the changes documented.

The growth of the vocal tract length in biological males is usually accounted for by the growth spurt during adolescence and the descent of the larynx. However, a change in FTMs' height or laryngeal descent is not reported in the studies so far. This suggests that there must be additional mechanisms beyond increasing height that can cause such an effect. Antoszewski et al.(2009) indicate that the size of female-to-male transsexuals' maxillary and mandibular canines and the first molars are between those of males and females. There is also anecdotal evidence based on

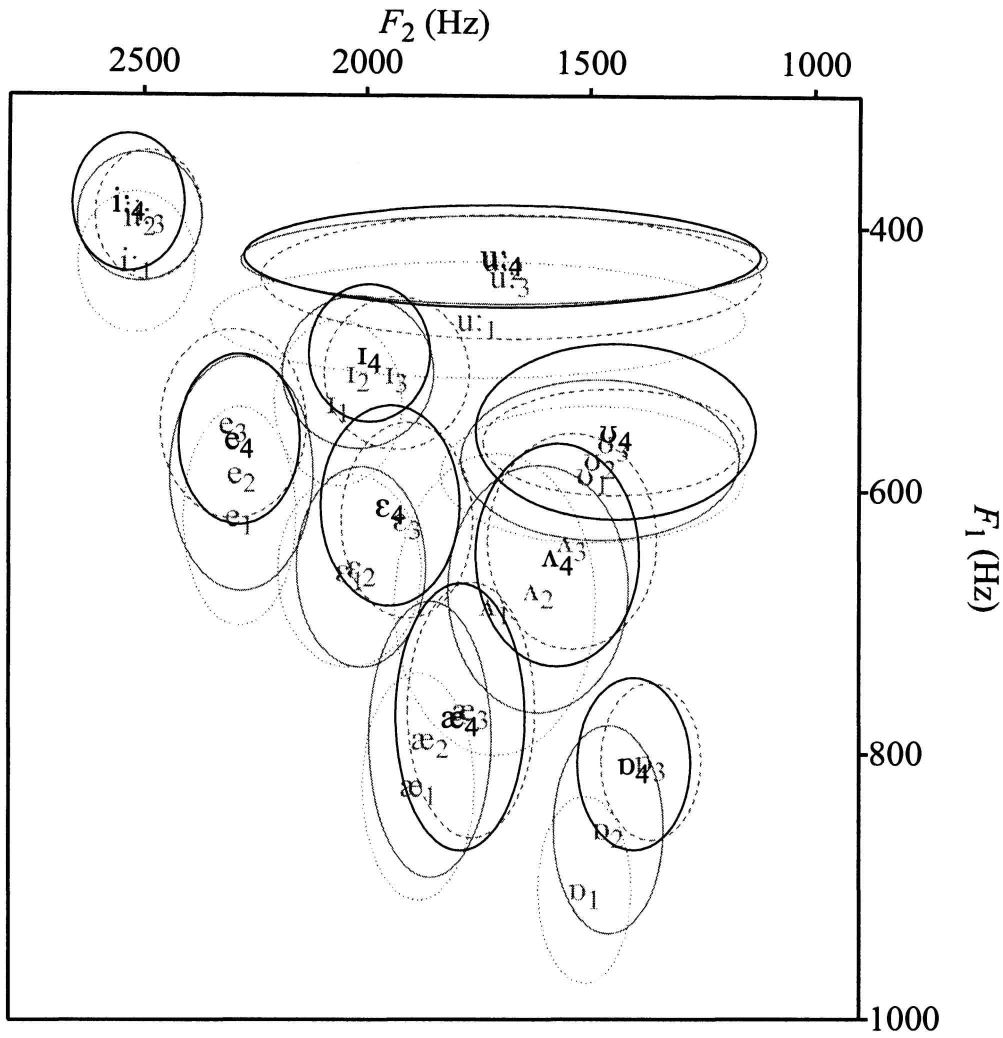

Figure 4.10 : F1 × F2 plot of CM over three years. Grey dotted line: pre-testosterone values; Grey solid line: values after 1 year on testosterone; Grey dashed line: 2 years; Black solid line: 3 years. The ellipses represent one standard deviation from the mean.

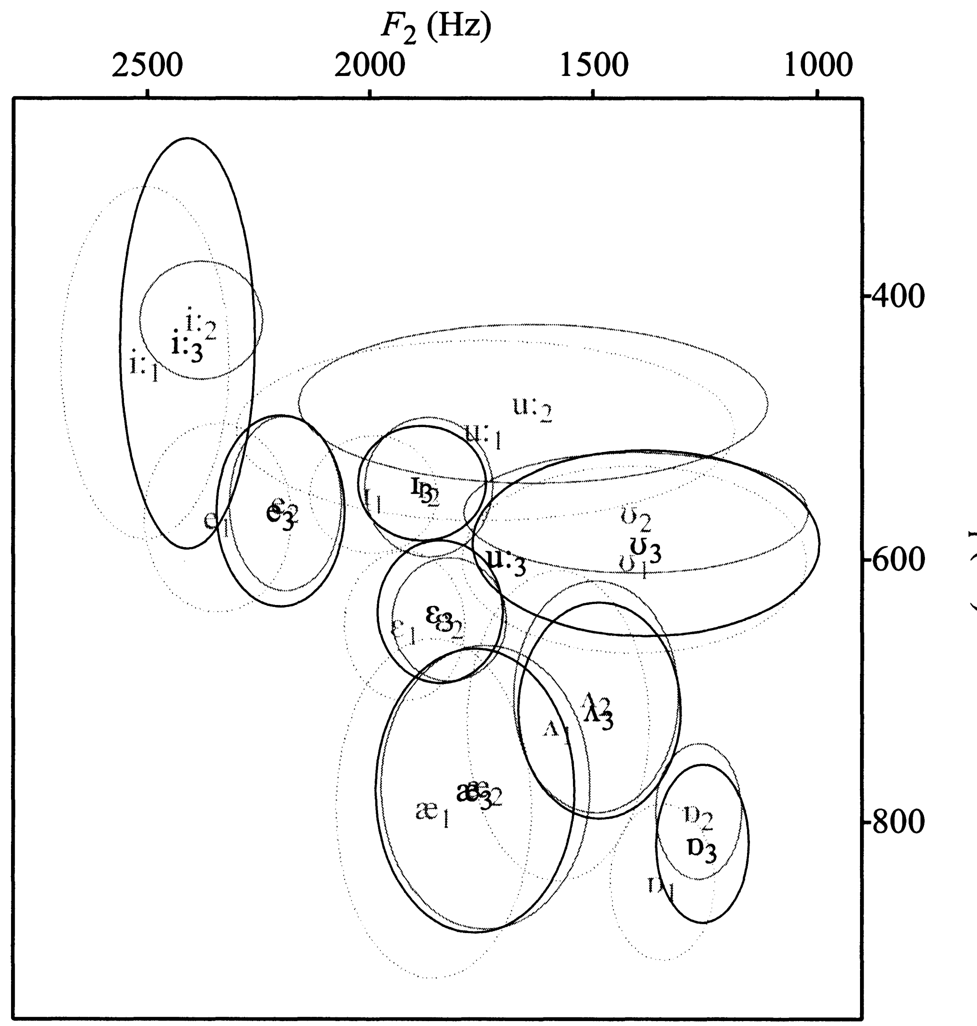

Figure 4.11 : F1 × F2 plot of LW over two years. Grey dotted line: pre-testosterone values; Grey solid line: values after 1 year on testosterone; Black solid line: 2 years. The ellipses represent one standard deviation from the mean.

introspection by FTMs that indicates that the width of the chin, as well as the angle between the posterior part of the mandible and the lower border of the ramus change noticeably during transitioning, creating a more masculine jawline. Some transmen experience the growth of the bony protrusion of an Adam's apple, which may be accompanied with other laryngeal cartilage shifts and growths. All these phenomena indicate that smaller bony and cartilaginous structures in and around the vocal tract may remain sensitive to androgens comparatively late in life.

The changes in the mandible that affect both the transverse and the sagittal dimensions may bring about an expansion of the pharynx that, in turn, leads to an increase in the supra-laryngeal vocal tract length. Thus, it may be that physiological changes alone account for the formant lowering, albeit of a novel type.

On the other hand, behavioural changes could also help explain the phenomena. Coleman (1983) hypothesizes that the difference between male and female formant values may be primarily due to the fact that females tend to have smaller physical dimensions with regard to the cavities of the head and neck than men. Hence, the frequencies at which the vocal tract resonates in females can be expected to be higher. Gunzburger (1995), however, proposes that differences in vowel formant frequency values between men and women are too great to be due to anatomical differences alone. She suggests that speakers modify their vocal characteristics to conform to feminine and masculine speech stereotypes. One example of this may be that females use more lip spreading and a more anterior tongue carriage during speech than males, as these are considered characteristic of females. Behavioural explanations of the changes documented above would predict that FTMs during their hormonal and societal transitioning are consciously or unconsciously unlearning female or feminine articulatory gestures.

If transitioning is considered a unidirectional change towards more male targets, these two approaches cannot account for the significant rise in formant values well into the transitioning process. As evidenced by both CM and LW, speakers may produce their entire vowel system at values that they already "left behind" once previously during transitioning. What could explain this slight but significant raising of the formant values?

It might be tempting to find a psychoacoustic explanation that aligns proprioception with production as follows. As FTM vocal sources and filter functions go through slow but radical changes, the voicing musculature needs to recalibrate itself for heavier vocal fold mass, possibly longer vocal folds and vocal tract lengths. As a result, up until the calibration is done, the muscles might be acting under previous motor programming. As a result, they are more prone to obey the physical forces imposed by the new weights and new lengths in the vocal tract. Assuming that the speakers still retain their old frequency values in memory, once they have increased muscular control, they might revert to(wards) pre-testosterone perceptual targets. As a result, they slowly re-modify their current articulatory targets.

However, previous speech adaptation and auditory feedback manipulation studies have demonstrated that talkers spontaneously compensate for real-time shifts by altering their production in a manner opposite to perturbation introduced in the feedback. These studies suggest that auditory feedback is part of a control system that actively influences the accuracy of articulator movement. What these studies agree on is that this adaptation process is incredibly fast but also transient, and it washes out of the production system shortly after the manipulated feedback stops. These lingering after-effects or adaptations to the perceptual manipulation indicate that a representation is involved in the control scheme. Compensation and adaptation have been demonstrated for formants (Houde & Jordan, 1998, 2002; Jones & Munhall, 2003; MacDonald, Goldberg, & Munhall, 2010; Munhall, MacDonald, Byrne, & Johnsrude, 2009; Purcell & Munhall, 2006; Villacorta, Perkell, & Guenther, 2007), fundamental frequency (Burnett, Freedland, Larson, & Hain, 1998; Jones & Munhall, 2002, 2005; Liu & Larson, 2007), and fricative centroids (Shiller, Sato, Gracco, & Baum, 2009) with this paradigm.

MacDonald et al. (2010) observes that the similarity of the results in their studies across experiments suggests that the compensatory response is dependent on the magnitude of the perturbation but not on the rate at which the perturbation is introduced. Liu and Larson (2007) also note that while the function relating compensation to formant-shift magnitude was linear for small perturbations, it was non-linear and "compressive" for large perturbations similar to results for pitch perturbations. Additional finding in Munhall et al. (2009) was that despite explicit instruction to ignore the feedback changes, subjects produced a robust compensation in all conditions, even when told to ignore the headphone or explicitly instructed to resist compensation. Their results suggest that compensation in the face of formant perturbation is relatively automatic, and the response is not easily modified by conscious strategy.

The changes in transmen's formant values are gradual, subtle and persistent. The gradualness predict that representations are very rapidly changed to reflect newer articulatory configurations. The persistence with which the changes happen also do not allow for a "washout" period during which the adapted values could be rolled back to their old values.

Alternatively, a behavioural explanation can be outlined as follows. As subjects experience the virilizing effects of testosterone on their whole bodies, such as changing fat-muscle ratio and distribution, increased muscle power, facial hair, male patterned hair, altered body odour and chemistry, the fatigue-combating effects of testosterone, etc., the relative importance of having an indubitably male identified voice to "pass" may become somewhat diminished. As a result, subjects' performance shifts from values that are consciously and actively (over-)masculinized to something that is perhaps less masculine, less physiologically predictable, but performs their current sex or gender identity better.

Because the data presented here are from two case studies, it is not possible to generalize for all individuals undergoing female to male gender reassignment. An investigation of the impact of androgen therapy on the filter function of the female larynx should be undertaken on further subjects to more accurately quantify and characterize the changes to formant values, and their relationship to changes in fundamental frequency, as well as the timing of these changes. Further investigation should be conducted with imaging methods to support the acoustic analysis and to situate the resulting formant values in male and female population statistic data.

Chapter 5

Acoustic correlates of transmen's gender and sexual orientation in sentence-level read speech

5.1 Abstract

This chapter presents a cross-sectional acoustic analysis of individuals who identify as transmen. Although the acoustic manifestations of sexual orientation in speech and voice have been researched, only a handful of studies address acoustic correlates of sexual orientation and trans history disclosure status in transgender individuals. So far the sole acoustic study on this population has only documented the differences between the researcher-established categories of gay-sounding cismales, and straight-sounding cismales vs.transmen (Zimman, 2010). The current study adds to the extremely sparse body of literature concerning the acoustic properties of transmen of various sexual orientation and it is the first to consider not only the sexual orientation but the trans history disclosure status of transmen as variables shaping the acoustic signal. In addition, the current study is the first foray into considering transmen's production in context of a larger data set of dialectally relevant sample, the Houston Urban English Survey (HUES) database (Niedzielski & Koops, 2011).

5.2 Sexual orientation of transmen

Before physically or socially transitioning, many (but certainly not all) transmen who are attracted to females identify, for lack of more flexible terminology, as lesbian. After physically transitioning, many FTMs consider themselves men and no longer identify as transsexuals. As a result, despite having been assigned female at birth, FTMs having male identities and attractions towards women may label their sexual orientation as heterosexual. Similarly, FTMs who are attracted to males may identify as gay post-transitioning. As Meier (in-press) points out, earlier studies on this population fluctuate in terminology depending on whether the authors defined sexual orientation based on birth sex or societal gender.

The prevalence of homosexuality in trans populations is a research topic that has been long ignored based on the heteronormative assumption that the reason behind transitioning is to "become straight." In Meier's survey, 605 self-identified FTMs from 19 different countries completed an online survey assessing their sexual orientation, sexual orientation identity, depression, anxiety, stress (Depression Anxiety Stress Scales), social support (Multidimensional Scale of Perceived Social Support), and health related quality of life (SF-36v2 Health Survey). Over half the sample (52%) reported attractions to both men and women. Meier's work makes a clear case for a more flexible system of sexual orientation classification for FTMs, as 35% of the sample reported a shift in sexual attractions during or as a result of transitioning.

Meier's findings also suggest that gay FTMs that are attracted to birth-assigned men ("heterosexual" transmen according to sex assigned at birth) are more prevalent than initially understood. To further complicate the current organization of sexual orientation, labels and partner preferences among transsexuals may change after hormone therapy or gender affirming surgery.

Prior linguistic research has overlooked both cisgender and transgender people whose sexual orientation is often not as simple as "heterosexual" or "homosexual." No peer-reviewed research has examined bisexual speakers, and most studies grouped these individuals with people of same-sex attractions (e.g., Munson, McDonald, DeBoe, & White, 2006; Zimman, 2010). Meier's work, however, provides a reason for considering FTMs who are attracted to both men and women separately. He found that bisexual FTMs may experience different levels of depression, anxiety, stress, support from their communities and families, and quality of life compared to FTMs who are attracted to either only men or only women. Meier's review of the psychosexuality literature draws on recent research on the non-transgender LGB population which highlighted the importance of separating persons who identified as bisexual from those who identified as gay. For example, Herek et al. (2009) found bisexual men to be more likely to have negative attitudes and self-stigma related to their sexual orientation than gay or lesbian participants. These considerations justify using a three-way differentiation of sexual orientation in this study that considers societal gender instead of birth sex.

5.3 Sexual orientation and speech

5.3.1 Anecdotal accounts

A common folk linguistic belief holds that many members of the LGB community use a distinctive speech style that allows naive listeners to identify the speakers' sexual orientation even in the absence of visual clues and an overt disclosure. Anecdotal reports characterize gay men's speech style as having a wide pitch range and strongly fluctuating intonation, lisping, or a higher-pitched voice (see reviews in Gaudio (1994)). Descriptions and anecdotal recollections of lesbian women's speech patterns are less common, but generally portay more monotone pitch patterns than heterosexual women (see reviews in Moonwoman-Baird (1997) and Waksler (2001)). However, the use of similar descriptors by transgender speakers, and the success of those descriptors in correctly suggesting the speaker's sexual orientation to listeners, have not been explored systematically.

5.3.2 Experimental accounts

As Munson et al. (2006) pointed out, one of the major reasons why LGB speech is difficult to approach with discrete tools is that, much like research on other more canonical stylistic variants, this research focuses on a social group that is not only less visible than most, but is culturally and socio-economically highly diverse with only so many members of the community opting into the stylistic choices. This heterogeneity admittedly calls the generalizability of the experimental results into question. Zimman (2010), in his overview of the literature, pointed out controversial findings where some of the studies confirmed stereotypical speech features, some contradicted them, and yet others found no significant differences between LGB and straight speakers' production. Zimman's explanation, based on Zwicky (1997) is that the findings of previous studies defy integration into a coherent model because multiple styles can be interpreted as gay-sounding by the basis of differing from straight-sounding norm(s).

While the existence of gay male stylistic variants has been largely ascertained both in production and listeners' perception (Gaudio, 1994; Munson et al., 2006), specifying the acoustic features cueing the gay female style has not been as successful. Waksler (2001) examined pitch range in San Francisco Bay Area women's read speech (n = 24), and found no differences between self-identified heterosexual and L/B women. Pierrehumbert et al. (2004) found that LGB people on average produced more peripheral vowels compared to their same-sex heterosexual peers. For G/B men, this was due to an overall hyperarticulation of the vowel space (hypothesized to be the acoustic evidence for overly clear speech) and for L/B women this manifested in lower and backer articulation for /ɑ/ and /u/. The rest of the vowels were produced at comparable formant frequency values in both the L/B and straight female samples.

Among the variables whose production was investigated in previous research were vowel formant values (Pierrehumbert et al., 2004), pitch range (Gaudio, 1994; Smyth, Jacobs, & Rogers, 2003; Waksler, 2001), fricative centroids and duration (Crist, 1997; Munson et al., 2006; Smyth et al., 2003), /l/-fronting and aspiration (Smyth et al., 2003).

5.3.3 Fricatives

Previous research has shown that the acoustic characteristics of /s/ and /ʃ/ differ between gay / bi and heterosexual men. Indeed, variation in the properties of /s/ has been consistently shown to correlate with the perception of men's voices as gay- or straight-sounding. Durational variation of fricatives (Linville, 1998; Smyth & Rogers, 2002), as well as spectral moments, such as center of gravity, skewness and kurtosis (e.g., Linville, 1998; Munson et al., 2006; Munson & Babel, 2007) were investigated in the cisgender population.

Zimman's (2010) study compared three FTMs to two gay-sounding males and three straight-sounding males. The eight speakers were hand-picked by Zimman by auditory perception only (as opposed to self-labeling), a choice corroborated by eight other pilot listeners. His fricative results based on 20 tokens of /s/ and one token of /ʃ/ indicated that the FTM group had a significantly higher center of gravity in the energy distribution of /s/ than either the gay-sounding males or straight-sounding males. Therefore, I posit the following hypothesis:

Hypothesis 1. *Based on the spectral quality of female and gay fricative production, it is hypothesized that transmen that are attracted to males will exhibit higher center of gravity than other transmen (1a), more negative skewness (1b) and lower kurtosis (1c) values in both /s/ and /ʃ/. There will be no difference between the center of gravity (1d), skewness (1e) and kurtosis (1f) values of transmen that are attracted to both males and females and transmen that are attracted to males. There will be no effect of sexual orientation on the duration of the fricatives (1g).*

5.3.4 Pitch

Gaudio (1994) claimed that his gay-sounding male speakers employed greater pitch variation than the straight-sounding controls: they used more of the pitch range and changed pitch more frequently. However, this difference was found in only one of the two speaking tasks (the accounting reading passage), and of the 13 measures, only one showed a statistically significant difference, and eight approached significance.

Thus, I predict the following:

Hypothesis 2. *Transmen who are attracted to males will exhibit larger sentential pitch range (as expressed in semitones) than transmen who are attracted to females (2a). There will be no difference between the sentential pitch range of transmen who are attracted to both males and females and transmen who are attracted to males (2b).*

Subsequent studies that have investigated mean fundamental frequency (e.g., Linville, 1998; Smyth & Rogers, 2002; Smyth et al., 2003) found no statistically significant difference for this measure between gay and bisexual(-sounding) cismales. However, Munson et al. (2006) analyzed words produced in isolation rather than connected speech, and found that the gay-sounding men in their study in fact had higher fundamental frequency than the straight-sounding men.

In Zimman (2010), neither mean F_0, nor F_0 range correlated with speaker group, gayness rating, or masculinity rating. The variation within his three groups of three speakers was greater than the variation across the three "orientation" groups.

Hypothesis 3. *Finally, it is hypothesized that transmen who are attracted to males will exhibit higher mean sentential pitch than transmen who are attracted to females*

(3a). There will be no difference between the sentential mean pitch of transmen who are attracted to both males and females and transmen who are attracted to males (3b).

5.3.5 Vowel formants

Studies that compared mean F1 and F2 across gay- and straight-sounding speakers showed no significant overall differences (Linville, 1998; Munson et al., 2006; Smyth & Rogers, 2002; Pierrehumbert et al., 2004). Some studies, however, such as Munson and Babel (2007), found significant differences in individual vowels and / or formants. As an example, Munson and Babel (2007) found that gay-sounding cismales had higher mean F1 than straight-sounding cismales.

Zimman's (2010) results were based on the mean F1 and F2 calculated across 11 stressed vowels (3 × /ae/, 2 × /ɛ/, 3 × /i/, 2 × /a/ and 1 × /ʌ/). The results showed no significant differences in the sexual orientation groups' overall mean first and second formants. A few of the individual formants in individual lexical items (e.g., F2 in the word "trapped") did produce significant differences, however. The extremely small number of tokens and speakers makes these results questionable.

In this study, therefore, I expect that:

Hypothesis 4. *There will be no formant difference between transmen according to their sexual orientation. That is, transmen who are attracted to males, females or both will have similar F1-F3 values.*

5.4 (Non-)Disclosure of trans history

5.4.1 Closet vs.stealth

Similar to the LGB community, members of the trans community often need to make conscious choices about "being out" or "coming out" to friends, co-workers, church members, etc. While the life experiences and identities of the trans community are widely varied, the core meaning of the concept "stealth" means the non-disclosure of trans history.

Determining when or whether to employ or break stealth hinges on the capability to pass or present as the intended gender, as well as the settings and participants in a particular situation. In the absence of a perceived gender transgression, it is the transperson's choice if and when they disclose personal trans history. Recent work focusing on disclosure by transpeople sheds light on two important ways this differs from disclosure of sexual orientation of LGB people.

As Zimman (2009) observed in his ethnographic work, there are two distinct ways a person can come out as transgender: before and after a change in societal gender. The differences between coming out before and after transitioning presents a challenge to previous characterizations of coming out. For people who completed their transitioning from one gender role to the other, and whose gender identity is consistent with how they present, as judged by naive interlocutors, coming out does not mean revealing a gender identity, but revealing gender history of the individual shifting from one gender category to another. In order to distinguish the two types, Zimman uses "declaration" to refer to the "initial claiming of a transgender identity," and "disclosure" to refer to "sharing ones transgender history after transition" (p. 60). As a result of this, disclosure of a transgender past may undermine what the speaker sees as their true identity rather than illuminating it.

Schilt's work (2004, 2010) based on sociological interviews with transgender / transsexual males (FTMs) also found that disclosure in the two "phases" of a transperson's life does make an important difference in how the individual sees the act of disclosure. In Schilt's study, individuals who openly transitioned from female to male and remained in the same workplace, viewed self-disclosure of transgender identity as both a step toward positive mental health and an educational tool for co-workers and management. Individuals who made the transition and then went stealth, saw disclosing at work as not only unnecessary but also a breach of workplace etiquette. This latter group of stealth-at-work FTMs views disclosure not as becoming or displaying their "true selves" as transmen but rather as a loss of their true gender identity as males.

Besides the psycho-societal consequences of disclosure, being open or going stealth also has an impact on an individual's socio-economic status. Studies on LGB populations, such as the one by Ellis and Riggle (1995), found that openly LGB workers report higher job satisfaction. However, these findings are dampened by their second finding: workers who do not disclose make more money than openly LGB workers, which suggests that being open requires making economic sacrifices.

Schilt (2004) admits it is arguable whether or not any person can really be or display their "true" self at their workplace. However, she found the same prevailing attitude in the majority of autobiographies, self-help books, and scholarly work about transsexuals / transgender individuals up until the mid 1990s. Therapists and doctors recommended that transsexuals remain in their current jobs until the effects of hormones became too noticeable and then to quit, transition, and reenter the workforce as their newly affirmed gender. This practice, however, meant keeping their history as transsexuals hidden.

5.4.2 Hansbury's psycho-social model of transmasculine identities

Hansbury (2005) introduced a taxonomy of transmasculine identities based on his own experience as a self-identified transman and observer of the community. He arranged these identities on a continuum from a male-identified / essentialist position ("Woodworkers") to a less binary / more constructivist perspective ("Genderqueers"), with "Transmen" in the middle of the transmasculine continuum*.

In Hansbury's continuum, the Transman is someone who presents to the world as a man, but unlike the Woodworker, is more likely to let some portion of the world know about his transsexuality and female past and celebrate it to an extent. According to

*It is important to emphasize that Hansbury's research participants often claimed multiple labels that occupy different locations on the continuum. That is, his participants located themselves along the continuum depending on context. He describes the Woodworkers as "your run-of-the-mill, Joe-six-pack, female-to-male transsexuals" who live as men, are out to very few people, and keep their female histories hidden. He notes that they tend to be the older members in the FTM community. Hansbury attributes this to a pre-1990s, pre-gender-studies ideology in which blending into the woodwork was the primary goal of sexual reassignment.

Hansbury, the Transman's dilemma is "to be or not be invisible." The Genderqueer identity, Hansbury continues, defies classification in the gender binary. In the case of those who use no hormones, physically there is no difference between a "genderqueer boi" and a "butch dyke," as it is a matter of self-interpretation and definition.

Thus, based on the psychological, social and economic need of stealth transmen to keep their current male gender identity intact and their female gender history hidden, it is anticipated that stealth transmen will display more masculine overall values. This leads to the following hypothesis:

Hypothesis 5. *Stealth transmen will have lower F_0 (5a), lower sentential pitch range (5b) and lower formant (5c) values than non-stealth transmen (Hansbury's 'Transmen' and 'Genderqueers').*

Studies on male-female differences in fricative production (e.g., Jongman, Wayland, & Wong, 2000) reported that adult female productions of /s/ and /ʃ/ generally tended to have higher frequency spectra than productions by male speakers of similar age. Specifically, Jongman et al. (2000) found that female fricative productions exhibited significantly higher values for spectral peak location, spectral mean, and kurtosis, as well as lower measures of skewness. This leads to the next hypothesis:

Hypothesis 6. *Stealth transmen will display lower fricative center of gravity (6a), lower kurtosis (6b) and more positive skewness (6c) than non-stealth transmen.*

5.5 Gendered speech production

In the research on male-female differences in production, a wealth of explanations have been offered for the established differences, such as appealing to social (Eckert & McGonnell-Ginet, 1999; Gordon, 1997), cultural (van Bezooijen, 1995), anatomical / physiological (A. P. Simpson, 2001, 2002), and perceptual (Diehl, Lindblom, Hoemeke, & Fahey, 1996) factors. For an overview of these approaches see Chapter 2.

While the majority of the work on gender focused on cisgender speakers, since the 1990s there has been an increasing number of studies devoted to the exploration of transgender speech. For an overview of these see Chapter 3 and Chapter 4. Besides grappling with statistical and reliability issues due to the very small number of subjects in each of the studies, work on TG populations noticeably lacks integration of the findings with the results from larger data sets of cisgender speech production. In order to provide a background against which FTMs' data can be interpreted, in this study FTM formant data are compared to the results of the Houston Urban English Survey (Niedzielski & Koops, 2011).

Research from the 1950s through the 1990s (e.g., Peterson & Barney, 1952; Fitch & Giedd, 1999) repeatedly found statistically and perceptually significant male-female differences in vowel formant production. As females tend to be shorter than males within a population, adult female vocal tracts tend to be shorter than those of adult males. Accordingly, female formants tend to be higher in frequency. Based solely on the physics behind the tube model of vowel production, this would predict that if a female's vocal tract is about 15% shorter than a male's, then her formant frequencies would be uniformly higher by about 15% relative to his. However, as Fant (1975) showed, the scaling factor between female and male formant values is strongly non-uniform across either vowel categories or individual formants.

As the mean of transmen height closely approximates that of average cisfemales, their vocal tracts are also expected to be closer to female length. Since FTMs need to present as male members of the society, it is likely that transmen will try to modify the filter function to approximate cismale values. In this endeavour, testosterone seems to help them (cf. Chapter 4). Thus, I formulate the following hypothesis:

Hypothesis 7. *Transmen relative to cisfemales will have lower formant values as measured through the first three formants (F1-F3) (7a). Transmen's formants will remain higher than cismale values (7b).*

5.6 Participants

Participants consisted of 35 female to male transgender talkers. Talkers were recruited from the members of Houston STAG (Some Transgenders Are Guys) self-help group, the Transgender Foundation of America (TFA) and from the greater Houston metropolitan area. Emails were circulated on the STAG listserv and in the community to advertise a study examining the FTM voice. No reference to sexual orientation, stealth status or hormone history was made in the email flyer.

The participants were between 18 and 51 years of age. Most of the participants were from the greater Houston metropolitan area, many of them undergraduates and graduate students in the Houston tertiary educational system, with a few out-of-town visitors.

Participants reported no history of language disorders. Two participants reported stuttering and miscellaneous speech disorders (cluttering and lisping), one participant reported dyslexia, and one hearing loss in one ear. All reported being native speakers of English, and three reported being bilingual (two Spanish-English bilinguals and one Malay-English bilingual). Participants were predominantly from the southern United States, with some speakers from other states. Most of them reported high mobility due to their trans status. This implies that this group of speakers displays significant variance in their regional dialect, which might have affected results. All of the participants had graduated high school, and were either enrolled in or had completed an undergraduate university degree. All but one participant had normal or corrected-to-normal vision. Participants were not compensated for completing the experiment.

At this point, it is worth repeating that using the term "transgender" in this study is intentional. I am using this term to denote the social nature of living as a man by participants who were assigned female at birth, regardless of whether they have taken steps to transition medically, surgically and legally or not. This means that subjects could enroll regardless of whether they had started testosterone, or not. As a result, participants ranged from zero to 156 months on testosterone. While in the medical, legal and psychiatric literature many other definitions of "transgender" exist, I wish to foreground the social as opposed to the strictly physiological aspects of transmen's speech.

5.7 Materials and stimuli

Stimuli for this experiment were the following six short sentences:

Pam, buy the dip.

Sam, try the dish.

Give some Silk to Sherry.

Give the milk to Kerry.

Philip, combine the malt with the fig.

Phyllis, combine the salt with the fish.

The stimuli sentences all contain at least one instance of the vowel /ɪ/. This vowel diphthongizes relatively infrequently across dialects (Labov et al., 2005) and it is close to a corner vowel, which makes it more suitable to map out the vowel space maximum in speakers than central medium height vowels would be. The six stimuli sentences come in three pairs. One member of each pair contains the fricatives /s/ or /ʃ/, while the other member is devoid of them.

All of the sentences deal with food and cooking, and it was therefore assumed that the vocabulary would be familiar to most participants. One vocabulary item, *Silk*, a brand of soy milk, was regularly read correctly in pilot tests (it being a short and not infrequent common noun as well), but some of the pilot testers, as well as some of the participants, asked verification questions about the meaning of the word in this context.

5.8 Procedures

During this task, the researcher presented the participants with a list of sentences in which the words were typed in 14-point Times New Roman font. One quasi-randomized list was created. Each participant read through the list twice. In order to ensure that there were at least two clear readings of each sentence, participants were asked to re-read any sentence they stumbled on, usually flanked by the previous and following sentence in the list.

Participants were asked to imagine that they were the narrators of a cooking video game and that they were giving simple orders to players in the game. Participants were not told that the stimuli would be used in a production experiment until the post-experiment debriefing. This task was embedded in a larger protocol consisting of four speech-production tasks, two of which are discussed in this dissertation.

The data were recorded in the laboratory sound-attenuated booth of the Department of Communication Sciences and Disorders at the University of Houston directly onto computer through a Kay Pentax 4600 interface with a AKG-C460 head-mounted unidirectional condenser microphone. The recordings were made at a sampling rate of 44.1 kHz, with 16-bit quantization.

In a post-recording questionnaire transmen were also asked to provide their height (in inches), the length of time they had been on testosterone (converted to months), and multiple choice questions about sexual orientation, gender labels the participants identified with, and their disclosure status.

5.9 Analysis

The acoustic measures were carried out using the Praat signal-processing program (Boersma & Weenink, 2010) with custom made scripts written by the author. The onset and offset of each vowel and fricative phoneme in each word was marked on an interval tier in Praat. The segmentation process was done by the author who was blind to the talker's demograpic survey data (including age, sexual orientation, disclosure status, time on testosterone, etc.). Both the manual segmentation and the acoustic analysis were done prior to and separate from the data entry and coding of the paper surveys, which was done by research assistants in psychology at the University of Houston. The research assistants were blind to the goals of the experiment and they received credit in Research Methods in Psychology for their work. Prior to conducting acoustic analyses, tokens containing extraneous noise, disfluencies, or reading errors were removed. Less than 2% of the data was lost in the process.

5.9.1 Filter characteristics: vowel formant frequencies

The first three formants (F1-F3) of each /ɪ/ vowel (the lexical category KIT) were measured. The formants were extracted from LPC formant analyses with 11 coefficients taken during the middle fifth of the vowel's duration. Mistracked and missing formant values were re-measured by hand.

5.9.2 Source characteristics: fundamental frequency and its descriptors

Two measures of the voicing source were taken. The measures were made based on the entire length of the individual sentences, excluding the last two syllables, which often devolve into vocal fry. The first measure was average fundamental frequency (F_0) over the sentence. The second source measure was the fundamental frequency range over the sentence. This was calculated by expressing the difference between the lowest and highest F_0 in the sentence in semitones.

5.9.3 Sibilant fricative measure

The final set of measures taken in this experiment were of the fricatives /s/ and /ʃ/. Besides duration, in this study three spectral measures were taken. These included center of gravity, skewness (the third spectral moment around the mean) and kurtosis (the fourth moment around the mean). These were calculated from the entire available frequency range, from 0 Hz to 22050 Hz. Nittrouer (1995) and McFarland et al. (1996) have reported that spectral moments 1, 3, and 4 (mean, skewness, and kurtosis, respectively) sufficiently distinguish /s/ from /ʃ/ across male and female adult English speakers and different vowel contexts. Specifically, /ʃ/ was characterized by a lower spectral mean, positive skewness, and smaller kurtosis, indicating a slightly flatter spectrum. Additionally, older speakers showed a more negative skewness in the distribution of energy in /s/, which is linked to both female as well as gay-sounding male speakers.

5.9.4 Sexual orientation and disclosure status

For the purposes of this study, participants who in Question 6.4 selected "Stealth male" (on its own or with other options), as well as participants who selected only "Male" (and nothing else) were categorized as stealth. Everyone else was categorized as non-stealth.

Following Meier (in press), sexual orientation was categorized by attraction. In creating the sexual orientation subgroups, transmen who in Question 6.8 selected only females (including females and / or transwomen) were categorized as "attracted to women" (AW). Transmen who selected only males and / or transmen were categorized as "attracted to males" (AM). Any other combination was categorized as "attracted to both" (AB).

6.4. *What word(s) best describe your gender identity?* (check all that apply)
Male
Stealth Male (living as male but not disclosing trans history)
FTM / transman / transmasculine
Genderqueer / gender fluid / third gender
Other, please specify

6.8. *What people are you attracted to?* (check all that apply)
Males
Females
Transmen
Transwomen
Genderqueer / gender fluid / third gender people
Intersex people
Other, please specify

Table 5.1 : Excerpts from the survey pertaining to disclosure status (6.4) and sexual orientation (6.8).

5.9.5 Statistical analysis

In order to describe the demographic data, a two-way ANOVA was run with disclosure status (DS) and sexual orientation (SO) as between-subjects variables, and age and height as the dependent variables.

To evaluate the hypotheses concerning formant values, for each set of dependent measures, a two-factor analysis of variance (ANOVA) was used. In each ANOVA, talker disclosure status and sexual orientation were the between-subjects factors. Across these ANOVAs, Bonferroni corrections were made for multiple comparisons.

The two summary measures of the voicing source (mean fundamental frequency and sentential frequency range) were submitted to a two-way ANOVA examining the influence of the independent variables sexual orientation and disclosure status.

Finally, the eight summary measures of sibilant fricative spectra (center of gravity, skewness, and kurtosis, as well as the duration of the two fricatives) were submitted to a two-factor ANOVA examining the influence of independent variables disclosure status and sexual orientation.

Means and standard deviations for all dependent measures are given on page 126 in Table 5.2, grouped by talker sexual orientation and disclosure status. For a visual overview see Figure 5.1 and Figure 5.2.

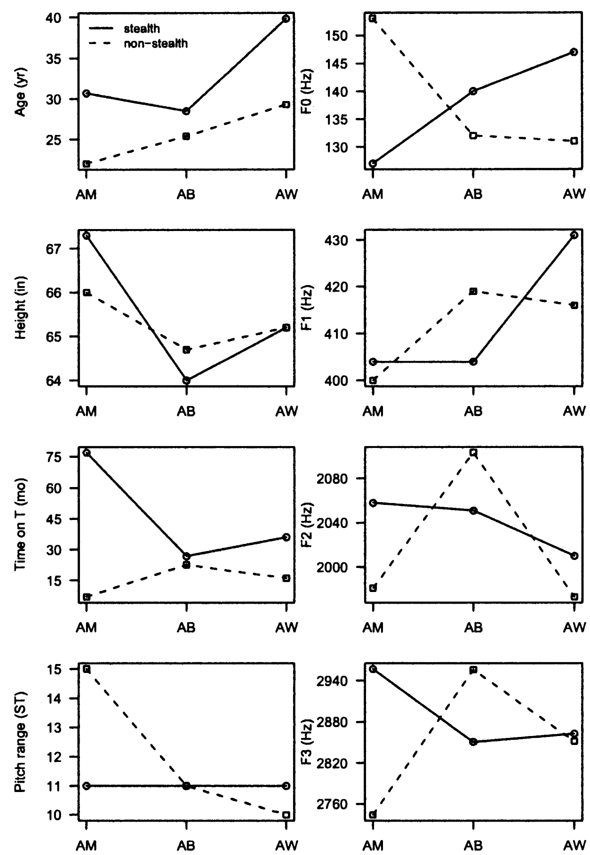

Figure 5.1 : Summary of demographic, fundamental and formant frequency descriptors in the cross-sectional study.

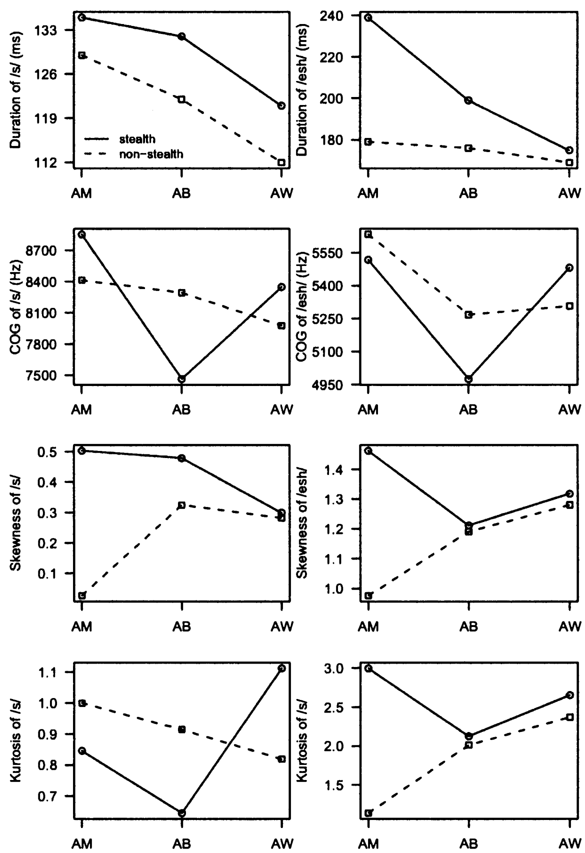

Figure 5.2 : Summary of durational and spectral descriptors of fricatives in the cross-sectional study.

		non-stealth				stealth			
		AB	AM	AW	Total	AB	AM	AW	Total
F_0 (Hz)	mean	132	153	131	133	140	127	147	141
	SD	29	10	20	25	46	10	22	28
F_0 range (ST)	mean	11	15	10	11	11	11	11	11
	SD	3	3	3	3	4	2	3	3
F1 (Hz)	mean	419	400	416	417	404	404	431	419
	SD	62	44	52	57	65	45	60	59
F2 (Hz)	mean	2103	1981	1973	2039	2051	2058	2010	2029
	SD	203	62	181	197	166	205	217	203
F3 (Hz)	mean	2956	2744	2852	2898	2851	2957	2863	2881
	SD	280	125	214	252	211	393	233	271
dur /s/ (ms)	mean	122	129	112	118	132	135	121	125
	SD	31	30	29	30	31	25	39	36
COG /s/ (Hz)	mean	8293	8414	7976	8155	7463	8854	8348	8163
	SD	1023	953	843	948	1127	606	986	1088
skewness /s/	mean	0.324	0.027	0.282	.2800	0.479	0.503	0.299	0.3688
	SD	0.303	0.313	0.287	.30505	0.195	0.352	0.316	.30440
kurtosis /s/	mean	0.915	1	0.819	0.8775	0.645	0.846	1.112	0.9587
	SD	0.75	0.783	0.734	0.74315	0.767	0.863	0.806	0.82245
dur /ʃ/ (ms)	mean	176	179	169	173	199	239	175	188
	SD	50	49	54	52	64	101	56	66
COG /ʃ/ (Hz)	mean	5267	5634	5307	5316	4975	5518	5482	5348
	SD	880	771	579	744	849	394	552	668
skewness /ʃ/	mean	1.191	0.976	1.28	1.215	1.212	1.463	1.318	1.304
	SD	0.34	0.092	0.313	0.323	0.285	0.198	0.234	0.253
kurtosis /ʃ/	mean	2.012	1.138	2.371	2.107	2.126	2.997	2.654	2.545
	SD	1.421	0.389	1.209	1.304	1.28	0.908	1.028	1.114
age (yr)	mean	25.4	22	29.3	26.9	28.5	30.7	39.8	35.3
	SD	6.2	-	6.2	6.2	8.3	2.1	10.3	10
height (in)	mean	64.7	66	65.2	65	64	67.3	65.2	65.3
	SD	2.6	-	2.4	2.4	2.9	1.5	2.6	2.6
time on T (mo)	mean	22.6	7	16.1	18.8	26.8	77	36.1	41.4
	SD	22.4	-	13.3	18.1	21.1	58.8	58.6	52

Table 5.2 : Summary of the acoustic descriptors in the cross-sectional study.

5.10 Results

5.10.1 Demographic information

The vast majority of the respondents identified as White / Caucasian. There was one African American, one West Asian American and one East Asian American in the sample. Out of the 34 subjects in this study, 18 were non-stealth and 16 were stealth. Seventeen men were attracted to females only, four to males only and 13 to both males and females. The average age was 30.8, with the minimum being 18 and the maximum 51.

It is noteworthy that all participants refrained from using the "Other, please specify" option in the questions about sexual orientation and disclosure status. Fifteen participants used a single option to describe their sexual orientation; the others used multiple terms. Table 5.3 on page 128 provides information on how participants labeled themselves. DS stands for disclosure status and '0' denotes non-stealth, while '1' denotes stealth. SO indicates the partner attraction of the participants. "AW" stands for "attracted to women," "AM" for "attracted to men," and "AB" for "attracted to both." The column "Hansbury" contains the taxonomy of the participants in Hansbury's (2005) tripartite system of Woodworker (W), Genderqueer (GW), and Transman (FTM).

Two separate two-way ANOVAs were run for age and height, respectively. The between subject variables were disclosure status (DS) and sexual orientation (SO). For age, there was a significant main effect of DS. Stealth FTMs in this sample were significantly older than non-stealth FTMs ($F\ (1,\ 34) = 4.397$, $p = 0.045$). Their average age was 35.3 ($SD = 9.9$), while the average age of non-stealth FTMs was 26.9 years ($SD = 6.2$). A significant main effect of sexual orientation was found as well ($F\ (1,\ 34) = 3.894$, $p = 0.032$). Bonferroni-corrected post hoc tests showed that FTMs who were attracted to women (AW) were significantly older than FTMs attracted to both men and women (AB) ($p = 0.017$). No other pairwise comparison was found significant. The AW group averaged 34.8 years ($SD = 9.9$), AM averaged 28.5 ($SD = 4.6$) and AB averaged 26.4 years ($SD = 6.7$). There was no significant interaction effect of disclosure status and sexual orientation.

For height, no significant main effect of sexual orientation or disclosure status was found. No significant interaction effect was found. For a summary of participant demographics see Table 5.4 and Table 5.5.

ID	DS	SO	Hansbury
1	1	AB	W FTM
2	0	AB	FTM GQ
3	1	AW	W
4	0	AB	GQ
5	0	AB	W FTM
6	1	AM	W
7	0	AW	FTM
8	0	AW	FTM
9	1	AB	W FTM
10	0	AM	FTM
11	1	AW	W
12	0	AW	FTM GQ
13	0	AB	FTM GQ
14	1	AW	W
15	1	AW	W
16	0	AW	FTM GQ
17	0	AB	FTM
18	0	AB	GQ
19	0	AW	FTM
20	0	AB	FTM
21	0	AW	FTM
22	1	AW	W FTM
23	0	AW	FTM
24	1	AW	W
25	1	AW	(missing)
26	1	AW	W
27	1	AW	W
28	1	AB	W FTM GQ
29	0	AB	FTM
30	1	AM	W
31	0	AB	FTM
32	0	AW	FTM
33	1	AB	W
34	1	AW	W
35	1	AM	W

Table 5.3 : Demography of participants.

	stealth	non-stealth	total
AB	4	9	13
AM	3	1	4
AW	9	8	17
total	16	18	34

Table 5.4 : Cross-tabular summary of sexual orientation and disclosure status.

	stealth	non-stealth	mean
AB	28.5	25.4	26.9
AM	30.7	22	26.3
AW	39.8	29.25	34.5
mean	35.2	26.9	30.8

Table 5.5 : Cross-tabular summary of age by sexual orientation and disclosure status.

5.10.2 Filter characteristics: vowel formant frequencies

The results of the ANOVA examining average F1, F2 and F3 showed no significant main effects of disclosure status, but F2 showed a significant main effect of orientation (F (2, 177) = 3.1, $p = 0.048$). Bonferroni-corrected multiple comparisons showed that AW had a significantly *lower* F2 in /ɪ/ than AB (AW: 1993 Hz, AB: 2088 Hz, AM: 2038 Hz; $p = 0.009$; Bonferroni-corrected alpha level = 0.016). No other comparisons were significant. Formant frequencies were in general lower in AM than in AW, with AB patterning either with AM or AW (see Table 5.6 on page 130). Disclosure status and sexual orientation did not interact significantly.

These results partially support Hypothesis 4, that there will be no formant differences between transmen according to their sexual orientation. Hypothesis 5b, that stealth transmen will have lower formant values than non-stealth transmen, was not supported.

F1 (Hz)		**AB**	**AM**	**AW**
non-stealth	mean	419	400	416
	SD	62	44	52
stealth	mean	404	404	431
	SD	65	45	60

F2 (Hz)		**AB**	**AM**	**AW**
non-stealth	mean	2103	1981	1973
	SD	203	62	181
stealth	mean	2051	2058	2010
	SD	166	205	217

F3 (Hz)		**AB**	**AM**	**AW**
non-stealth	mean	2956	2744	2852
	SD	280	125	214
stealth	mean	2851	2957	2863
	SD	211	393	233

Table 5.6 : Summary of formant values by sexual orientation and disclosure status (values rounded to the nearest Hz).

5.10.3 Source characteristics: fundamental frequency

The two summary measures of the voicing source (mean F_0, sentential F_0 range) were submitted to a two-way ANOVA examining the influence of sexual orientation and disclosure status. For an overview of group means see Table 5.7 on page 132.

For mean F_0 there was no significant main effect of sexual orientation or disclosure status. However, a significant interaction effect between sexual orientation and disclosure status (F (2, 172) = 4.982, p = 0.008) was found. AM transmen had lower mean sentential fundamental frequency when stealth and higher when non-stealth.

AB and AW transmen had lower fundamental frequency when non-stealth.

For sentential range there was a statistically significant main effect of sexual orientation (F (1, 172) = 4.211, p = 0.016) as well as a statistically significant main effect of disclosure status (F (1, 172) = 5.335, p = 0.022). Stealth transmen exhibited a smaller sentential range than other FTMs (10.9 vs.11.09 ST). Bonferroni-corrected multiple comparisons showed that AM employed a wider sentential range than did AB and AW (12.9 ST vs.10.9 and 10.7 ST, respectively, p = 0.009). Sexual orientation and disclosure status had a significant interaction on sentential range (F (2, 172) = 5.335, p = 0.006). AM transmen had an smaller sentential pitch range when stealth and a range 4 ST larger when non-stealth. Stealth AB transmen had a slightly smaller range than non-stealth AB transmen. Stealth AW transmen had slightly larger pitch ranges than non-stealth AW transmen. The average differences within AB and within AW were less than 1 ST.

Therefore, Hypothesis 2a, that AM transmen will exhibit larger sentential pitch range than AW transmen, was supported. However, Hypothesis 2b, that there will be no difference between the sentential pitch range of AM and AB transmen, was not supported. Hypothesis 3a, that AM transmen will exhibit higher mean sentential pitch than AW transmen, was not supported. Hypothesis 3b, that there will be no difference in the sentential mean pitch between AB transmen and AM transmen, was supported. Hypothesis 5a, that stealth transmen will have lower F_0 than non-stealth transmen, was not supported. However, Hypothesis 5b, that stealth transmen exhibit narrower sentential pitch range values than non-stealth transmen, was supported.

mean fundamental frequency (Hz)		**AB**	**AM**	**AW**
non-stealth	mean	132	153	131
	SD	29	10	20
stealth	mean	140	127	147
	SD	46	10	22

fundamental frequency range (ST)		**AB**	**AM**	**AW**
non-stealth	mean	11	15	10
	SD	3	3	3
stealth	mean	11	11	11
	SD	4	2	3

Table 5.7 : Summary of fundamental frequency descriptors by sexual orientation and disclosure status (values rounded to the nearest Hz and ST).

5.10.4 Sibilant fricative spectra

Eight summary measures of sibilant fricative spectra (center of gravity, skewness, and kurtosis, as well as the duration of the two fricatives) were submitted to a two-factor ANOVA examining the influence of disclosure status and sexual orientation.

The fricative /s/

For descriptive statistics of the fricative /s/ see Table 5.8 on page 133.

There was a significant main effect of disclosure status on the skewness of /s/ (F (1, 321) = 24.96, $p < 0.000$), with stealth transmen possessing a more positive skewness (0.369) than other FTMs (0.280).

Sexual orientation had a significant main effect on the duration (F (2, 326) = 5.434, $p = 0.005$), the center of gravity (COG) (F (2, 326) = 7.754, $p = 0.001$) and the skewness (F (2, 326) = 5.397, $p = 0.005$) of /s/. In duration, a Bonferroni-

duration (ms)		**AB**	**AW**	**AM**
non-stealth	mean	122	112	129
	SD	31	29	30
stealth	mean	132	121	135
	SD	31	39	25

COG (Hz)		**AB**	**AW**	**AM**
non-stealth	mean	8293	7976	8414
	SD	1023	843	953
stealth	mean	7463	8348	8854
	SD	1127	986	606

skewness		**AB**	**AW**	**AM**
non-stealth	mean	0.324	0.282	0.027
	SD	0.303	0.287	0.313
stealth	mean	0.479	0.299	0.503
	SD	0.195	0.316	0.352

kurtosis		**AB**	**AW**	**AM**
non-stealth	mean	0.915	0.819	1
	SD	0.75	0.734	0.783
stealth	mean	0.645	1.112	0.846
	SD	0.767	0.806	0.863

Table 5.8 : Summary of durational and spectral descriptors of /s/ by sexual orientation and disclosure status (values rounded to the nearest Hz and ms).

corrected multiple comparison showed no significant difference between the pairs. The duration of the fricative was 127 ms for AB, 116 ms for AW and 132 ms for AM (Table 5.9, page 134)†.

duration	group	average (ms)	SD (ms)
	AB	126	31
	AW	116	35
	AM	132	28

	group	significance
	AB-AW	0.056
	AB-AM	1.000
	AW-AM	0.052

Table 5.9 : Summary of duration group means and significances of Bonferroni corrected pairwise comparisons for /s/.

These results are similar in magnitude to the available data on gay-straight differences, such as the work of Linville (1998). Linville found that the duration of /s/ in gay speakers ranged between 105.8 and 119.7 ms (mean = 111.4 ms), while in straight speakers the duration ranged between 85.4 ms to 103 ms (mean = 95.1 ms). The direction of the difference is further attested in research on gay-sounding or gay-stereotyped voices (as opposed to the voice of self-identified gay speakers), such as Podesva, Roberts, and Campbell-Kibler (2002) and Smyth and Rogers (2002), as well as studies on the perception of resynthesized speech by Levon (2007). The size

†Even though there was a significant main effect on the duration, minor differences, such as the five ms difference between 127 ms and 132 ms, are plausibly within the margin of error in measurements. As such, this work makes no strong claims about these results. Further work is planned to establish within-rater reliability.

of the differences are also on par with the measured between-gender differences in English (onset) fricatives (Table 1 in Weglarski, Sewall, Schiavetti, Metz, and Whitehead (2000)) that, depending on the flanking vowel, ranged between 10 to 20 ms in /s/ and 20 to 30 ms in /s/.

As for center of gravity of /s/, the Bonferroni post hoc comparison showed a significant difference between AB and AM (7878 Hz vs.8634 Hz, $p = 0.006$, Bonferroni-corrected alpha level = 0.016) (see Table 5.10 on page 135).

COG	**group**	**average (Hz)**	**SD (Hz)**
	AB	7878	1123
	AW	8162	936
	AM	8634	816

	group	**significance**
	AB-AW	$p = 0.627$
	AB-AM	$p = 0.006$*
	AW-AM	$p = 0.045$

Table 5.10 : Summary of COG group means and significances for /s/; Bonferroni-corrected alpha level = 0.016 for a familywise error rate of 0.05. Significant comparison marked with asterisk.

In skewness measures, the Bonferroni post hoc comparison showed no significant difference between the pairs (see Table 5.11 on page 136).

Sexual orientation and disclosure status interacted significantly for /s/ center of gravity (F (2, 321) = 13.8, $p < 0.001$), skewness (F (2, 321) = 8.267, $p < 0.001$), and kurtosis (F (2, 321) = 4.73, $p = 0.010$).

The COG of /s/ of AM transmen was much higher when stealth than when non-stealth. The stealth COG value is somewhat higher than the non-stealth value in

skewness	group	average	SD
	AB	0.402	0.281
	AW	0.291	0.301
	AM	0.265	0.406

	group	significance
	AB-AM	$p = 0.209$
	AW-AB	$p = 0.052$
	AW-AM	$p = 1.000$

Table 5.11 : Summary of skewness group means and significances for /s/; Bonferroni-corrected alpha level = 0.016 for a familywise error rate of 0.05.

AW transmen. For AB transmen, the stealth COG value for /s/ was actually lower than the non-stealth. The skewness of /s/ of stealth AM transmen, again, was much higher than that of non-stealth transmen. AW transmen showed minimal difference across disclosure status, with stealth AW transmen having higher skewness than their non-stealth peers. Stealth AB transmen also had higher skewness values than non-stealth AB transmen, with the difference across disclosure status being between those of AW and AM transmen. Stealth AM and AB transmen had lower kurtosis values than their respective non-stealth peers. Stealth AW transmen, on the other hand, had higher kurtosis values for /s/ than did non-stealth AW transmen.

The fricative /esh/

For descriptive statistics of the fricative /ʃ/ see Table 5.12 on page 137.

There was a significant main effect of disclosure status on the duration (F (1, 196) = 7.426, p = 0.007), skewness (F (1, 196) = 11.318, p = 0.001) and kurtosis (F (1, 196) = 11.080, p = 0.001) of /ʃ/. Stealth transmen exhibited a longer fricative

duration (ms)		**AB**	**AW**	**AM**
non-stealth	mean	176	169	179
	SD	50	54	49
stealth	mean	199	175	239
	SD	64	56	101

COG (Hz)		**AB**	**AW**	**AM**
non-stealth	mean	5267	5307	5634
	SD	880	579	771
stealth	mean	4975	5482	5518
	SD	849	552	394

skewness		**AB**	**AW**	**AM**
non-stealth	mean	1.191	1.28	0.976
	SD	0.34	0.313	0.092
stealth	mean	1.212	1.318	1.463
	SD	0.285	0.234	0.198

kurtosis		**AB**	**AW**	**AM**
non-stealth	mean	2.012	2.371	1.138
	SD	1.421	1.209	0.389
stealth	mean	2.126	2.654	2.997
	SD	1.28	1.028	0.908

Table 5.12 : Summary of durational and spectral descriptors of /esh/ by sexual orientation and disclosure status (values rounded to the nearest Hz and ms).

than non-stealth transmen (188 ms vs.173 ms) with more positive skewness (1.304 vs.1.215) and higher kurtosis (2.55 vs.2.11).

There was a significant main effect of sexual orientation on the duration (F (1, 196) = 4.010, $p = 0.020$), center of gravity (F (1, 196) = 4.440, $p = 0.013$) and kurtosis (F (1, 196) = 3.203, $p = 0.043$) of /ʃ/. Bonferroni post hoc multiple comparisons showed no significant differences. For all the Bonferroni comparisons see Table 5.13 (page 139).

Sexual orientation and disclosure status interacted significantly for /ʃ/ skewness (F (2, 196) = 5.179, $p = 0.006$) and kurtosis (F (2, 196) = 5.609, $p = 0.021$). While the skewness values of /ʃ/ were very similar between AB and AW transmen regardless of disclosure status (between 1.191 and 1.318), stealth AM transmen had much higher skewness values than non-stealth AM transmen (0.976 vs.1.463). The stealth skewness values were higher than the non-stealth ones on all three sexual orientation groups. Similarly, the stealth kurtosis values were higher than the non-stealth ones, regardless of sexual orientation. The difference between stealth vs.non-stealth AM transmen was the largest, between stealth vs.non-stealth AB transmen the smallest, and the difference between stealth vs.non-stealth AW transmen was between them.

Based on the spectral quality of female and gay fricative production, transmen attracted to males were expected to exhibit higher center of gravity (1a), more negative skewness (1b) and lower kurtosis (1c) values. AM transmen indeed exhibited the highest center of gravity for both fricatives, but the Bonferroni-corrected t-tests reached statistical significance in /s/ only. AM transmen exhibited the most negative

duration	group	average (ms)	SD (ms)
	AB	188	56
	AW	172	55
	AM	209	82

	group	significance
	AB-AW	$p = 0.501$
	AB-AM	$p = 0.291$
	AW-AM	$p = 0.035$

COG	group	average (Hz)	SD (Hz)
	AB	5120	876
	AW	5394	569
	AM	5576	597

	group	significance
	AB-AW	$p = 0.101$
	AB-AM	$p = 0.085$
	AW-AM	$p = 0.950$

kurtosis	group	average	SD
	AB	2.069	1.368
	AW	2.512	1.122
	AM	2.067	1.171

	group	significance
	AB-AW	$p = 0.032$
	AB-AM	$p = 1.000$
	AW-AM	$p = 0.422$

Table 5.13 : Summary of group means and significances for /esh/; Bonferroni-corrected alpha level = 0.016 for a familywise error rate of 0.05.

skewness measure in /s/, but it did not reach statistical significance. AM transmen's kurtosis values did not show statistically significant differences from the other two groups. Thus, hypothesis 1a was partially supported, and 1b and 1c were not supported.

Hypothesis 1d, that there will be no difference between the center of gravity values of AM transmen and AB transmen, is partially supported: although there was a strong significant difference found in the center of gravity of /s/ between AM (higher COG) and AB transmen (lower COG), and while the difference in /ʃ/ is in the same direction, it did not reach significance. Hypotheses 1e-f, that there will be no difference between the skewness and kurtosis values of AM transmen and AB transmen were supported for both fricatives.

Hypothesis 1g, that there will be no effect of sexual orientation on the duration of the fricatives was supported. No significant differences were found in the two fricatives, even though the duration of /ʃ/ when produced by AM transmen was on average 40 ms longer than in the production of AW transmen.

Hypothesis 6a, that stealth transmen will display lower fricative centers of gravity than non-stealth transmen, was not supported in /s/. Hypothesis 6c, that stealth transmen will exhibit more positive skewness, was supported for both /s/ and /ʃ/. Hypothesis 6b, that stealth transmen exhibit lower kurtosis, was not supported for /s/. For /ʃ/, the opposite was actually found significant, namely, that stealth transmen exhibit higher kurtosis values than non-stealth transmen, was found significant.

5.11 FTM values in the context of the HUES database

Finally, Hypothesis 7 predicted that transmen relative to cisfemales will have lower formant values as measured through the first three formants (F1-F3) (7a) and that transmen's formants will remain higher than cismale values (7b). First, the appropriate formant values were queried in the Houston Urban English Survey database (Niedzielski & Koops, 2011), and then I pooled the results with the FTM database. The sex variable in HUES and the disclosure status variable in the FTM database were collapsed into a "gender identity" variable. Finally, one-way ANOVAs were run with gender identity of the subjects (male, female, stealth, non-stealth) as the between subject variable and F1, F2 and F3 as the dependent variables. For descriptive statistics of formant values by speaker group see Table 5.14 (page 141).

Formant (Hz)		**HUES male**	**HUES female**	**stealth FTM**	**non-stealth FTM**
F1	mean	421	516	419	417
	SD	53	82	59	57
F2	mean	1942	2333	2029	2039
	SD	205	200	203	197
F3	mean	2625	3049	2881	2898
	SD	184	150	271	252

Table 5.14 : Summary of formant means by speaker group.

Gender identity exhibited a significant main effect on all formants: F1 ($F\ (3, 1132) = 189.457$, $p < 0.001$), F2 ($F\ (3, 1132) = 306.916$, $p < 0.001$), and F3 ($F\ (3, 1132) = 417.411$, $p < 0.001$). Bonferonni-corrected post hoc tests showed that in F1 there was no significant difference between cismales (420 Hz), stealth (419 Hz) and non-stealth (416 Hz) transmen. Cisfemales (516 Hz) were significantly different from all three male groups ($p < 0.001$ in all three cases). For F2 the only pairwise comparison that was *not* significant was that between stealth (2029 Hz) and non-stealth (2039 Hz) transmen. Both groups of transmen were significantly higher than cismales (1942 Hz) and significantly lower than cisfemales (2333 Hz). Cismales and cisfemales had, not unexpectedly, significantly different F2 values in their /ɪ/ production, with cisfemales producing higher values. Similarly to the findings for F2, for F3 production only the paired comparison between stealth (2881 Hz) and non-stealth (2898 Hz) transmen was *not* significant. Both stealth and non-stealth transmen's production of F3 in the /ɪ/ vowel was significantly higher than that of cismales (2625 Hz), and lower than that of cisfemales (3049 Hz). For the Bonferroni-corrected pairwise comparisons see Table 5.15 on page 143.

In sum, Hypothesis 7a was supported for all three formants and 7b was partially supported, as F2 and F3 exhibited a statistically significant distinction between the two groups in the expected direction, but F1 showed no statistically significant difference between transmen's and cismen's production. For an F1 $\times$ F2 scatterplot of the findings see Figure 5.3 (page 144).

5.12 Discussion

The current study adds to the extremely sparse body of literature concerning the relationship between acoustic features of transmen's speech and their sexual orientation

F1	group	significance
	male-female	$p = < 0.001$*
	male-stealth	$p = 1$
	male-nonstealth	$p = 1$
	female-stealth	$p = < 0.001$*
	female-nonstealth	$p = < 0.001$*
	stealth-nonstealth	$p = 1$

F2	group	significance
	male-female	$p = < 0.001$*
	male-stealth	$p = 0.002$*
	male-nonstealth	$p = < 0.001$*
	female-stealth	$p = < 0.001$*
	female-nonstealth	$p = < 0.001$*
	stealth-nonstealth	$p = 1$

F3	group	significance
	male-female	$p = < 0.001$*
	male-stealth	$p = < 0.001$*
	male-nonstealth	$p = < 0.001$*
	female-stealth	$p = < 0.001$*
	female-nonstealth	$p = < 0.001$*
	stealth-nonstealth	$p = 1$

Table 5.15 : Summary of Bonferroni-corrected pairwise comparisons for formants F1-F3; Bonferroni-corrected alpha level = 0.0083 for a familywise error rate of 0.05. Significant comparisons marked with an asterisk.

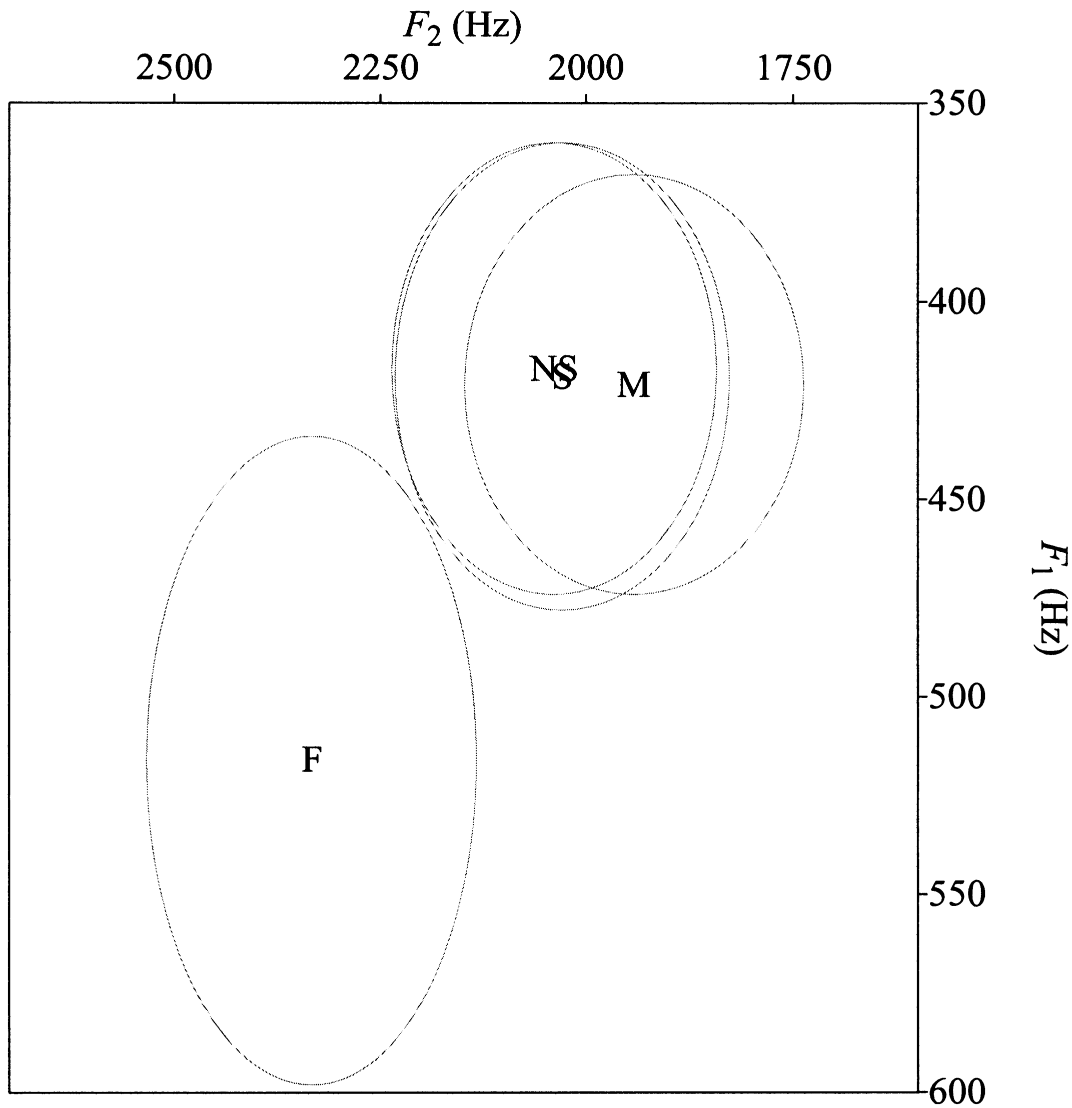

Figure 5.3 : F1 × F2 scatterplot of /ih/ vowel in the HUES and the FTM databases; F = HUES cisfemales, M = HUES cismales, S = stealth FTMs, NS = non-stealth FTMs.

and disclosure status.

The results reveal a number of general findings about transmen's speech. For instance, the sexual orientation of transmen had a significant effect on the F2 of /ɪ/, sentential pitch range, and fricative center of gravity. Sentential mean pitch, fricative duration, skewness and kurtosis differences were biased towards the expected direction but did not reach significance.

In sum, AW transmen produced the shortest fricatives and AM transmen the longest, with AB transmen between them. AM transmen had the highest center of gravity for both fricatives; AB the lowest with AW transmen between the two groups. Because a measure of center of gravity indicates the center of the distribution, it correlates negatively with the length of the front resonating cavity. Thus, a higher COG value indicates a shorter front cavity, or in articulatory terms, a fronted place of articulation. This is consistent with findings in the literature on female and gay male fricatives, both of which tend to have higher COGs. Together with the longer duration, these two variables build a case for the increased perceptual salience of these fricatives in the production of transmen who are attracted to males.

Jongman et al. (2000) and Zimman (2009) found that more negative skewness in the distribution of energy in /s/ is linked to both female as well as gay-sounding male speech. Consonant with this, AM transmen also exhibited low(er) spectral skewness values (lowest for /s/ and low for /ʃ/) compared to the AW and AB groups, whose spectral skewness values for /s/ were similar. This is also consistent with the findings of Munson et al. (2006) where the spectra of gay and bi men's /s/ was significantly more negatively skewed than that of heterosexual men.

The spectral kurtosis values indicate the peakedness of the spectral shape, and there is evidence that it may differentiate between apical and laminal articulatory gestures of the fricatives (Li et al. 2009). It is noted that females typically produce higher kurtosis values, that is, more sharply defined peaks (Jongman et al., 2000). However, the picture emerging from the spectral kurtosis is far from simple. The kurtosis value for /s/ exhibited by AM transmen is high and patterns with AW transmen, but the kurtosis value for /ʃ/ exhibited by AM transmen is low and patterns with AB transmen. For /ʃ/, the kurtosis value for AW transmen was the highest of the three groups.

In interpreting the results, it is important to bear in mind that there is ample evidence that /s/ varies substantially within individuals for reasons that have nothing to do with identity. Consider the study of Maniwa, Jongman, and Wade (2009) concerning the acoustics of fricatives spoken in conversational and intentionally clear-speech styles. They found that clear-speech tokens of /s/ were louder, longer, and had higher peak frequencies than those in conversational speech styles.

Similar to the findings of Pierrehumbert et al. (2004) and Rogers, Smyth, Jacobs, Jensen, and van Herk (2003), who found that hyperarticulated vowel spaces were associated with G/B men's speech, AM and AB transmen in this study had significantly higher F2 values than AW transmen did. The finding that AM transmen employed a wider sentential range than did AB and AW transmen is consistent with Gaudio's (1994) results suggesting that more gay-sounding men produced speech with greater pitch ranges than less gay-sounding men. My results also validated the unorthodox methodological choice to investigate transmen attracted to men separately from transmen who are attracted to both men and women.

Turning to disclosure status, fewer differences in the speech of stealth transmen and non-stealth transmen emerged than were expected based on the literature and transmen's anecdotal experiences. Stealth and non-stealth transmen had very similar formant values, sentential fundamental frequency, and fricative center of gravity, but they differed significantly in the sentential pitch range and had more positive fricative skewness, kurtosis, and fricative duration.

This investigation differs from previous studies on gender, sexual orientation, disclosure status and speech in multiple ways. First, it examined the speech production of transmen of various sexual orientations during various phases of transitioning using a consistent set of speech materials and acoustic analyses. This highlights the need for additional work that deals not only with the differences — the "otherness" of these talkers from normative, non-stigmatized (heterosexual, cis-gendered) or less-stigmatized (homosexual, cis-gendered) talkers — but the variation within this group.

The second way in which this study differs from previous research is that the speech materials used in acoustic and perceptual analyses were sentences constructed for their phonetic composition. That is, this investigation differs from other studies (e.g., Munson et al., 2006) and is similar to others (Pierrehumbert et al., 2004; Rogers et al., 2003; Smyth et al., 2003) in that it examined sentence-level stimuli instead of single-vowel or single-word productions.

Since connected speech is the basis of everyday communication, the measures from these stimuli have the advantage of phonological and phonetic richness that word- or phone-level stimuli cannot compete with. Production tasks based on sentences also boost the naturalness of the task over synthetic or unnaturally short stimuli to base judgments on. Read speech has markedly different phonetic quality from spontaneous speech, but the present study uses read speech due to the research trade-off between control over variables and the naturalness of the stimuli. Admittedly, connected speech brings with it a number of complications, as Munson et al. (2006) pointed out. For instance, voiceless stop consonants have longer voice onset times at the left edge of a prosodic phrase, vowels and consonants are hyper-articulated when they occur in pitch-accented syllables, and prosodic variation leads to phonetic variation in vowels.

Because of the prosody vs.segmental phonetic interaction, as Munson et al. (2006) point out, in regression analyses the explanatory weight of segmental cues can be potentially masked by intonational cues co-occurring or carrying said segmental cues. Munson et al. (2006) bring forward two earlier studies that found that the linguistically more complex the stimuli were, the more accurately listeners identified sociolinguistically relevant categories (race in Lass, Tecca, Mancuso, and Black (1979) and regional dialect in van Bezooijen and Gooskens (1999)). However, since one of the aims of the current studies is to quantify how successfully transmen pass as males in absence of visual clues (e.g.over the phone or at a drive-through), sentence-level stimuli have the advantage of simulating real-life encounters.

This study did not attempt to document the existence of FTM speech styles (or lack thereof) or whether they conform to popular cultural stereotypes. What it attempted, however, was to demonstrate how linguistic resources are shared in this FTM population. The theoretical background behind "Darsey's theorem" (Kulick, 2000, p. 259) and the so-called "all and only test" (Wong, Roberts, & Campbell-Kibler, 2002) state that research on gays and lesbians need to establish that (a) gays and lesbians use language in empirically delineable ways, and (b) those ways of using language are unique to all gays and lesbians and no one else. While the second condition is intuitively impossible, my findings support the results of current research, namely, that no social group ever pass the "all and only test". To use Queen's (2002, p. 73) succinct terminology, both the research focus in this field and the interpretation of the results should move on the continuum of relative "sameness-in-difference" and "difference-in-sameness."

Similar to the findings of Podesva et al. (2002) (also published as Podesva, Roberts, and Campbell-Kibler (2006), my findings problematize this notion of a single gay or single straight style or way of speaking. I demonstrated that not all members of any categories in this study deployed pitch, formant or fricative based resources consistently, even though these have been linked to gay, straight, female, etc. identity in the literature. I also demonstrated that every group shared these resources. This is consistent with the way Livia (2002) models variation: groups emerge not by the presence or absence of a variable bundle, but by the degree and combination of these deployed resources.

Finally, a quote from Eckert may illuminate why AM transmen seemed to frequently strongly differ from other transmen: "I have shown elsewhere that the most extreme users of phonological variables in my adolescent data are those who have to do the greatest amount of symbolic work to affirm their membership in groups or communities" (Eckert, 1989, p. 59). Popular belief (which is sadly attested in the DSM-IV as well) still holds that transsexuals transition into a straight-identified body. Therefore it is clearly AM transmen whose double outlaw status (transsexual *and* gay) forces them to constantly exhibit who they are, as who they are is defined with respect to majority males (straight) and (possibly imagined) majority transsexuals (also straight)[‡].

5.13 Future research directions

The results from this study suggest further lines of inquiry into the transmasculine experience. Besides investigating FTMs of colour and older (mostly stealth) FTMs, an inquiry into different levels of disclosure looks promising. One approach would be to evaluate FTMs' production along a Transgender Outness Inventory based on the LGB Outness Inventory by Mohr and Fassinger (2000). The Outness Inventory was originally created for lesbians and gay men to assess level of outness to different people in different spheres of functioning, especially in workplace relations. For trans people, level of outness about gender identity as well as level of outness about assigned birth sex should be explored.

Zimman's (2009) distinction between the processes of declaration and disclosure,

[‡]cf. the findings of Meier (in-press)

the "initial claiming of a transgender identity," vs. "sharing one's transgender history after transition," created a tool that ties FTMs' production to another important construct: "passing." Passing in the gender variant and drag community means to successfully present as the intended gender.

Examining the interaction, correlation and / or causality, of passing and disclosure may shed further light on the linguistic "gender security" that can prompt transmen to deploy or break stealth. Examining prominent transman Jamison Green's autobiographical work *Becoming a Visible Man*, Raj (2011) points out romanticised "plasticity of gender" in Butler's post-Gender-Trouble works. In Raj's opinion it introduces a new essentialism, namely, that crossing is privileged over passing, which in turn could lead to categorizing transgender (and transsexual) bodies as subversive or normative, depending on whether they cross or pass. However, the so far anecdotal, and only recently research-borne knowledge (e.g., Meier, in-press) of the interaction between age of commencement of testosterone therapy, disclosure status and partner attraction seem to suggest that there is a social drift in the FTM community from straight male-identified stealth men towards openly queer and openly gender-queer or gender-variant individuals. Investigating how the crossing-passing-disclosure triad is constructed through language is a very exciting area of future research in this field.

Kulick (2000) appeals to the notion of "transitivity" when making a case for approaching language from desire as opposed to from sexuality. Sexuality, he says, tends to be viewed as what "one has" or "one is" and thus the research focuses on the subject: how they reveal or conceal their sexuality. Desire, however, is definitionally transitive in the sense that it is always directed toward someone or something. An additional advantage of desire is that it is not specific to any particular kind of person. "What are specific to different kinds of people are the precise things they desire and the manner in which particular desires are signalled in culturally codified ways" Kulick (2000, p. 273).[§] In light of this, the results of this present study, or the methodological approach in future studies, could be reconceptualized along desires. One such approach would be reformulating the already mentioned dichotomy between "attracted to gender" vs. "attracted to bodies" as the desire to be intimate with similarly oriented vs.similarly configured individuals.

[§]Critiques of Kulick's approach (e.g., Queen, 2002, p. 73) point out that focusing on desire as opposed to sexuality does not entail escaping categorization as it only starts a different, new process of categorization, which inevitably leads to navigating between essentialism and constructionism.

Chapter 6

Summary and outlook

6.1 Triple decoupling

In the Butlerian view of performativity, it is talk that produces the talker and not the other way around. Results of the studies on the speech production of FTMs present phonetic evidence that transmen's speech produce the transmen by, following and adapting Bagemihl's terminology (1997), what I term triple decoupling.

Transmen successfully decouple gender from biological sex. The results from the longitudinal studies exemplify that if you are born female and passed puberty as a female, you do not necessarily need to have a female voicing source or filter function. Both qualitative changes can be achieved (albeit to different degree) by bringing exogenous testosterone into the system that will in turn virilize both the source and the filter over time. Moreover, the cross-sectional study showed an additional phenomenon, namely that the formants of the /ɪ/ vowel patterned closer with cismale values than cisfemale values. Admittedly, the entire vowel space was not analysed, so we cannot ascertain if the rest of the vowels show a similar pattern. However, the abrupt difference between the production of cisgender females in the HUES database and the production of transmen, as well as the proximity of cismales and transmen, indicate that articulatory gestures can be modified to move the acoustic targets towards a gendered target one is striving to present.

The acoustic manifestations of the groups of transmen with different partner attraction offers the next type of decoupling, that between sexual orientation and gender identity. The results of the cross-sectional study imply that if you were born female and are attracted to men, you do not necessarily have to identify as a woman. You can opt in / out of the self-identification by selectively adopting features associated with the gay cismale speaking style. Sexual orientation was found to have a significant main effect on the duration, the center of gravity, and the skewness of /s/, as well as the duration, center of gravity, and kurtosis of /ʃ/. Similarly to the findings of Gaudio (1994), Pierrehumbert et al. (2004), and Rogers et al. (2003), evidence of hyperarticulated vowel spaces as exemplified by the significantly higher F2 values of AW transmen and wider sentential range was found in the production of transmen who were not attracted to women or women only.

Lastly, the longitudinal studies provide evidence for the third type of decoupling, that which comes in the form of gender breaking free from physiology. The recurring "reverse J-pattern" of both the transitioning source and filter, as well as the mean fundamental frequency hovering above the pitch floor illustrate the fact that transmen do not need to feel obliged to sound as masculine (as low-pitched and "low-formanted") as testosterone enables them to. This final type of decoupling also serves to demonstrate that many transmen decidedly do not want to opt in to the binary system of sex / gender – even though they could. Thus, the linguistic production of transmen successfully challenges those older sociolinguistic approaches (often referred to as the "coat-rack" model) that still focus on the sociolinguistic manifestations of a (pre-performance, pre-discoursive) biological sex which maps nearly perfectly over social gender.

In social and philosophical theory the notion "agency" refers to the capacity of individuals to act independently and to make their own free choices (Ahearn, 2001). "Structure," by contrast, refers to the recurrent patterned arrangements which seem to influence or limit the choices and opportunities that individuals possess. While in linguistic anthropology, agency is typically located in discourse, my work shows how agency can be attested at the level of the segment and fundamental frequency descriptors, even in scripted and non-discoursive speech. The three decouplings listed above are ways of transmen successfully exhibiting agency.

6.2 Limitations of the dissertation

An integral part of research methodology is recognizing the biases of the research and how they might influence the findings. Biases may arise from the demographic characteristics of the sample, of the researcher as well as the researcher's values and beliefs about the topic being studied. There were several biases that influenced the way data were collected, understood and interpreted in this dissertation. This section summarizes them and the efforts I made to mitigate them.

6.2.1 Sampling

This study relied on a non-random sample of transgender individuals who may have been more willing to record and disclose than a truly random sample of transgender individuals. The sample characteristics were also limited in this study as there was virtually no representation of people of colour or older transmen. There was also a skew in level of education in that many of the participants had advanced degrees and almost all of them had at least some college education. This was likely due to the recruitment process, as people were primarily recruited through snowball and other "friend-of-a-friend" convenience sampling methods.

I did not attempt the targeted recruitment of older transmen and transmen of colour, which might have significantly biased the sample to include, e.g.more straight stealth people. As both older stealth transmen and people of colour may define masculinity, femininity and transmasculinity differently than a younger population of mostly white transmen, further research is needed for the results to be even remotely generalizable to the wider transmasculine population.

The difficulty in finding transmen even in a huge, fairly trans-friendly metropolitan area like Houston contributed to the relatively small sample size, which can admittedly pose problems with reaching statistically significant and reliable results. As research becomes more prevalent in this area, I am hopeful that it may become easier to record data from the trans population.

6.2.2 Bins and boxes

A second limitation of this study involves the categorization schema of both sexual orientation and disclosure status. Having to "bin" participants' choices along two malleable continuums required imposition of a structure and creation of "boxes," even though the boxes broadened the options for representing identities. Specifically, because of the small number of participants, I had to collapse the two non-stealth categories that Hansbury (2005) called "transmen" and "genderqueers" and treat them as one overarching group of people. With regard to sexual orientation, I chose to define attraction by the gender of the people my participants were attracted to. While this choice was not arbitrary, "binning" participants by the "bodies" they are attracted to would have redrawn the groups significantly. For example, according to the system used in this study, transmen attracted to cismales and transwomen (be the latter pre-, post- or non-operative, on or off hormones) were categorized as "attracted to both." However, in an alternative system they could very well be categorized as people "attracted to (partially) male bodied individuals."

Even though this study worked with a small number of participants, on a relatively small number of tokens per variables, the findings amply highlight the breadth of diversity within the transmasculine community in Houston. As was exemplified by both the results of sexual orientation and disclosure status, the findings support an existing constructivist to essentialist continuum proposed by Hansbury (2005) as a useful way to organize transmasculine identities. However, it is important to emphasize that the participants in this study often claimed multiple labels that occupy different (sometimes non-adjacent) locations on the continuum. That is, participants located themselves along the continuum depending on context and participants (p.c. participants during de-briefing). Some of the participants specified that they are stealth at work but in the larger LGBTQ community they are "out" as drag-performing genderqueer transmen. This suggests further degrees of orientational and disclosure fluidity (as evidenced by Meier, in press) which merits caution in generalizing the results to other FTM communities, or even the same community in a few years' time.

Transgender and transsexual individuals themselves embody process and fluidity at multiple levels. This makes research on transgender language challenging by virtue of the independent variables leaching out of and across the analyst's categories. As a result, attempting to firmly anchor acoustic features by connecting them with transgender bodies, gender, socialization, partner attraction or disclosure status runs the risk that the analysis inadvertently reiterates binary, essentialist views of these bodies or practices. Luckily, and ironically, it seems that transgender and transsexual bodies, since they are often (partially or fully) incongruous with any essence that they are supposed to represent, are themselves the safeguards against being grossly binarized. Naturally, if this work still exhibits signs of essentialist values, it is my fault only.

6.2.3 Methodological bias

The quantitative approach to complex variables such as the ones referenced in my studies can tell us only so much about an identity. If I adopt the stance that identity is discursive and is performed, the few questions in the survey about orientation and disclosure status along with the handful of acoustic descriptors collected from scripted speech provide but a few minor facets of that performance, the representativeness of which can be easily questioned. Admittedly, a qualitative, ethnographic enquiry into the social practice of these individuals is likely a more adequate way to gain a better overall view of the identities in question. However, maintaining both the temporal resolution of the recorded material and my own needs for reliable and valid results necessitated a level of control over the variables that would not have been feasible for a single researcher or in a qualitative research framework.

This dissertation centered around the task of identifying and pointing out the linguistic resources that are used to index different partner attraction and disclosure status among transmen. Future investigations on this population could build on this is work and map out how these resources are used in discourse to index different meanings. For that, the adoption of a (more) qualitative approach is necessary that has its foundation in the analysis of discourse and cross-situational interactions.

6.2.4 Researcher bias

I am a cisgender female. As a result, I did not have the perspective or introspective experiences that would have enriched the data gathering and analysing process, particularly in the sense of recognizing the biasing factors due to my own cisgender privilege*. In order to challenge my own cisgender "outsider" beliefs, to recruit participants and to prove my commitment, in the past four years I actively participated

*Cisgender privilege refers to the "set of unearned advantages that individuals who identify as the gender they were assigned at birth accrue solely due to having a cisgender identity" (Walls & Costello, 2010, p. 83). For a check-list of such privileges see Taylor (2010).

in the life of the trans community in Houston and its satellite organizations. I volunteered and participated at social events put on by the Transgender Foundation of America, held and took workshops for transmen and transwomen, attended support group meetings, conferences, drag shows and ally trainings, collaborated with trans academics and participated in trans-oriented research as a cisgender control subject.

Problematizing the intersection of researcher bias and methodological bias, (Livia, 2002, p. 94–95) points out that while the theory has moved from "gender as indexed by sex" to theories of "gender as performative," the methodology has not changed much since. Researchers still identify the linguistic community, record them, segment the signal, measure and analyse things. However, Livia warns that the more researchers move away from older theories to new ones, the more they need to take their own relationship with the subjects into account. The question whether my subjects read me as a gay, straight, bisexual, female, androgynous, cisgender, transitioning, a friend, a community volunteer, or a researcher do merit caution in the interpretation of the results. The more researcher gender is seen as an influence on the stylistic arsenal of the subjects, the more possible meanings there are.

6.2.5 Terminological bias

Complicating the interpretation of the data is the fact that two individuals may identify with the same gender terminology, yet have very different lived experiences. This clearly promotes the use of open-ended questions rather than multiple choice questions to let participants define themselves. However, both open and close ended questions assume that the researcher's and subject's connotations of a given term are the same. There is no perfect solution to this problem, because while allowing participants to describe their identities is the least restrictive way of gathering data, in a quantitative study the researcher will often have to turn complex, continuum-like variables into categorical variables to be able to group participants.

6.3 Merits

This dissertation aimed to contribute to the gathering momentum of academic work on identities previously viewed as deviant, non-normative and often pathologized. I hope that it succeeded in moving transgenderism, transsexuality, and specifically female-to-male views thereof even further on their way towards the mainstream of scholarly discussions.

My approach attempted to not only provide a descriptive account of the acoustic features of transmen's speech, but to also emphasize the existence and the features of within-gender variation together with the between-gender similarities on par with the observable between-gender variation in the context of the HUES database.

6.4 Applications

One of the outcomes of the focus on social-indexical variation is a growing realization that the linguistic resources which have typically been considered in variationist studies (e.g., formant frequencies, fricative centroid, /t/ glottalization, etc.) are not the effects but the source of identity. Relating to the "bins and boxes" problem already mentioned above, is the realization that the social categories are imposed by the researcher since there is nothing intrinsic to those social variables that lead speakers to perform in a particular way. For example, being a young, educated gay male from Houston does not in itself determine that he would have a certain percentage of /s/ tokens with high fricative centroid. Instead sociophoneticians are now tapping into a much more complex process in which speakers deploy linguistic resources to position themselves within the communities they interact with. This is evidenced by style-shifting that speakers employ in different contexts, driven by an implicit convergence with (or divergence from) other conversational participants. This sort of style shifting or code-switching can to an extent be expected from transmen as well, who in certain situations must perform as one particular (gender) identity (e.g.stealth workplace presentation) and in other situations as other gendered identities (e.g.within the transgender community, at the gynaecologist's office, etc.).

Socially constructed beliefs center around factors such as perceived prestige and perceived aesthetic qualities of a particular variant. These beliefs further shape both laypeople's and professionals' behaviour and lead to the situation where some forms are highly salient (e.g., the "gay lisp"), whereas others are abundantly present but with far lower overt awareness on the part of speakers and listeners (e.g., vowel shifts). In the case of SLPs, the extent to which personal and medical ideologies relate to sociophonetic variation could influence clinical assessment and therapy. To exemplify this, Docherty and Khattab (2008) bring the example of clinical assessment of labial and labiodental variants of /r/ in speakers of British English, the voice quality of Liverpool English speakers, and Munson (2010) reviews the hybrid variationist-clinical study by van Borsel et al. (2009) that implicitly pathologizes the "gay lisp."

The point that ties together these studies is that a clinician's interpretation of what is normal or not for an individual speaker with regard to linguistic resources needs to be mediated by an understanding that those resources are aspects of the individual's speech production that contribute to defining and performing the individual's identity. This is very succinctly illustrated in the interview with prominent "gender outlaw" Kate Bornstein: "At voice lessons I was taught to speak in a very high-pitched, very breathy, very sing-song voice and to tag questions onto the end of each sentence. And I was supposed to smile all the time when I was talking. And I said, 'Oh, I don't want to talk like that!' The teachers assumed that you were going to be a heterosexual woman. No one was going to teach you to be a lesbian because lesbian was as big an outlaw as transsexual" (Bell, 1993, p. 112).

In the transgender community, the ability to pass as the desired gender is considered success. As already outlined in the longitudinal chapters, my experiences in the FTM community show that there are transmen who clearly need help from SLPs. These include transmen exhibiting the effects of persistent vocal fatigue, the inability to project their voice at gender-appropriate fundamental frequencies, the temporary or permanent loss of singing voice, and the transmen whose habitual speaking pitch "outs" them as non-normative males, or in worse cases, as females (even in the face of visually available secondary sexual characteristics).

In my longitudinal studies three out of the six transmen had problems passing well into their time on testosterone when their predicted optimal pitch, but not their habitual pitch, was already unambiguously masculine. This calls for assessments by SLPs and / or ENTs and, if necessary, speech therapy designed for transmen that simultaneously focuses on mean and maximum fundamental frequency descriptors, voice quality and breath management, and possibly formant values, either as part of the regular general medical assessment for transmen transitioning or as necessary therapy helping transitioning. Whether the therapy should be thought of as abandoning, shifting or creating other available styles for the transman to (code-)switch between is defined by the needs dictated by the identity of the transman.

References

Abu-Al-makarem, A., & Petrosino, L. (2007). Reading and spontaneous speaking fundamental frequency of young Arabic men for Arabic and English languages: A comparative study. *Perceptual and Motor Skills*, *105*(2), 572–580.

Adler, R. K., Hirsch, S., & Mordaunt, M. (Eds.). (2006). *Voice and communication therapy for the transgender/transsexual client: A comprehensive clinical guide.* Plural Publishing.

Adler, R. K., & van Borsel, J. (2006). Female-to-male considerations. In R. K. Adler, S. Hirsch, & M. Mordaunt (Eds.), *Voice and communication therapy for the transgender/transsexual client: A comprehensive clinical guide* (pp. 139–168). Plural Publishing.

Ahearn, L. M. (2001). Language and agency. *Annual Review of Anthropology*, *30*, 109-37.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders (DSM-IV-tr)* (4th ed.). Washington, DC: American Psychiatric Association.

Andrews, M. L., & Schmidt, C. P. (1997). Gender presentation: Perceptual and acoustical analyses of voice. *Journal of Voice*, *11*(3), 307–313.

Antoszewski, B., Zadzinska, E., & Foczpanski, J. (2009). The metric features of teeth in female-to-male transsexuals. *Archives of Sexual Behavior*, *38*(3), 351-358.

Bagemihl, B. (1997). Surrogate phonology and transsexual faggotry: A linguistic analogy for uncoupling sexual orientation from gender identity. In A. Livia & K. Hall (Eds.), *Queerly phrased: Language, gender, and sexuality* (pp. 380–401). New York: Oxford University Press.

Baker, J. (1999). A report on alterations to the speaking and singing voices of four women following hormonal therapy with virilizing agents. *Journal of Voice*, *13*(4), 496–507.

Barbieri, R. L., Evans, S., & Kistner, R. W. (1982). Danazol in the treatment of endometriosis: Analysis of 100 cases with a 4-year follow-up. *Fertility and Sterility*, *37*(6), 737–746.

Beck, J. M. (1999). Organic variation of the vocal apparatus. In W. J. Hardcastle & J. Laver (Eds.), *The handbook of phonetic sciences* (pp. 256–98). Blackwell Publishing.

Bell, S. (1993). Kate Bornstein: A transgender transsexual postmodern Tiresias. In A. Kroker & M. Kroker (Eds.), *The last sex: Feminism and outlaw bodies.* New York: St. Martin's Press.

Beranek, L. L. (1949). *Acoustic measurements* (1st ed.). John Wiley & Sons.

Bevis, R. R., Hayles, A. B., Isaacson, R. J., & Sather, A. H. (1977). Facial growth response to human growth hormone in hypopituitary dwarfs. *The Angle Orthodontist*, *47*(3), 193–205.

Boersma, P., & Weenink, D. (2010). *Praat 5.2.22.*

Boone, D. R., McFarlane, S. C., Berg, S. L. V., & Zraick, R. I. (2009). *The voice and voice therapy* (8th ed.). Allyn & Bacon.

Boothroyd, C. V., & Lepre, F. (1990). Permanent voice change resulting from Danazol therapy. *The Australian and New Zealand Journal of Obstetrics and Gynaecology*, *30*(3), 275–276.

Braun, A. (1995). Fundamental frequency – how speaker-specific is it? , 9–23.

Britto, A. I., & Doyle, P. C. (1990). A comparison of habitual and derived optimal voice fundamental frequency values in normal young adult speakers. *Journal of Speech and Hearing Disorders*, *55*(3), 476–484.

Burnett, T. A., Freedland, M. B., Larson, C. R., & Hain, T. C. (1998). Voice F0 responses to manipulations in pitch feedback. *The Journal of the Acoustical Society of America*, *103*(6), 3153–3161.

Butler, J. (1999). *Gender trouble: Feminism and the subversion of identity.* Routledge.

Carew, L., Dacakis, G., & Oates, J. (2007). The effectiveness of oral resonance therapy on the perception of femininity of voice in male-to-female transsexuals. *Journal of Voice*, *21*(5), 591–603.

Case, J. L. (2002). *Clinical management of voice disorders* (4th ed.). Austin, TX: Pro-Ed.

Coleman, R., & Markham, I. W. (1991). Normal variations in habitual pitch. *Journal of Voice*, *5*(2), 173–177.

Coleman, R. O. (1983). Acoustic correlates of speaker sex identification: Implications for the transsexual voice. *The Journal of Sex Research*, *19*(3), 293–295.

Constansis, A. (2008). The changing female-to-male (FTM) voice. *Radical Musicology*, *3*.

Cooksey, J. M. (2000). Male adolescent transforming voices: Voice classification, vocal skill development and music literature selection. In L. Thurman & G. Welch (Eds.), *Bodymind and voice: Foundations of voice education* (pp. 811–841). Collegeville, MN: The VoiceCare Network.

Crist, S. (1997). Duration of onset consonants in gay male stereotyped speech. , *4*(3), 53–70.

Damrose, E. J. (2009). Quantifying the impact of androgen therapy on the female larynx. *Auris, Nasus, Larynx*, *36*(1), 110–112.

Davies, S., & Goldberg, J. M. (2007). Clinical aspects of transgender speech feminization and masculinization. In W. O. Bockting & J. M. Goldberg (Eds.), *Guidelines for transgender care* (pp. 167–97). Informa Health Care.

de Bruin, M. D., Coerts, M. J., & Greven, A. J. (2000). Speech therapy in the management of male-to-female transsexuals. *Folia Phoniatrica et Logopedica*, *52*, 220–7.

Diamond, L. M. (2003). What does sexual orientation orient? A biobehavioral model distinguishing romantic love and sexual desire. *Psychological Review*, *110*(1), 173–192.

Diamond, L. M., & Butterworth, M. (2008). Questioning gender and sexual identity: Dynamic links over time. *Sex Roles*, *59*(5-6), 365–376.

Diehl, R. L., Lindblom, B., Hoemeke, K. A., & Fahey, R. P. (1996). On explaining certain male-female differences in the phonetic realization of vowel categories. *Journal of Phonetics*, *24*(2), 187–208.

Docherty, G. J., & Khattab, G. (2008). Sociophonetics and clinical linguistics. In M. Ball, M. Perkins, N. Müller, & S. Howard (Eds.), *The handbook of clinical linguistics* (pp. 603–625). Oxford, UK: Blackwell Publishing.

Eckert, P. (1989). The whole woman: Sex and gender differences in variation. *Language Variation and Change*, *1*, 245–267.

Eckert, P., & McGonnell-Ginet, S. (1999). New generalizations and explanations in language and gender research. *Language in Society*, *28*(2), 185–201.

Eckert, P., & McGonnell-Ginet, S. (2003). *Language and gender.* Cambridge, UK: Cambridge University Press.

Ellis, A., & Riggle, E. (1995). The relation of job satisfaction and degree of openness about ones sexual orientation for lesbian and gay men. *Journal Of Homosexuality*, *30*(2), 75–85.

Fairbanks, G. (1960). *Voice and articulation drillbook* (2nd ed.). New York: Harper & Row.

Fant, G. (1975). *Non-uniform vowel normalization* (Nos. Vol. 2–3). Stockholm, Sweden.

Fitch, W. T., & Giedd, J. (1999). Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *The Journal of the Acoustical Society of America*, *106*(3), 15111522.

Garvin, H. M. (2008). Ossification of laryngeal structures as indicators of age. *Journal of Forensic Sciences*, *53*(5), 1023–1027.

Gaudio, R. P. (1994). Sounding gay: Pitch properties in the speech of gay and straight men. *American Speech*, *69*(1), 30–57.

Gauffin, J., & Hammarberg, B. (1992). *Vocal fold physiology: Acoustic, perceptual and physiological aspects of voice mechanisms.* Whurr Publishers.

Gelfer, M. P. (1989). Stability in phonational frequency range. *Journal of Communication Disorders*, *22*(3), 181–192.

Gelfer, M. P., & Schofield, K. J. (2000). Comparison of acoustic and perceptual measures of voice in male-to-female transsexuals perceived as female versus those perceived as male. *Journal of Voice*, *14*(1), 22–33.

Gerritsma, E. J., Brocaar, M. P., Hakkesteegt, M. M., & Birkenhger, J. C. (1994). Virilization of the voice in post-menopausal women due to the anabolic steroid nandrolone decanoate (Decadurabolin). The effects of medication for one year. *Clinical Otolaryngology and Allied Sciences*, *19*(1), 79–84.

Glasberg, B. R., & Moore, B. C. (1990). Derivation of auditory filter shapes from notched-noise data. *Hearing Research*, *47*(1-2), 103–138.

Gordon, E. (1997). Sex, speech, and stereotypes: Why women use prestige speech forms more than men. *Language in Society*, *26*(1), 47–63.

Gorton, R. N., Buth, J., & Spade, D. (2005). *Medical therapy and health maintenance for transgender men: A guide for health care providers.* San Fransisco: Lyon-Martin Women's Health Services.

Greene, M. C. L., & Mathieson, L. (1989). *The voice and its disorders* (5th ed.). Whurr Publishers.

Günzburger, D. (1989). Voice adaptation by transsexuals. *Clinical Linguistics & Phonetics*, *3*(2), 163–172.

Günzburger, D. (1993). An acoustic analysis and some perceptual data concerning voice change in male-female trans-sexuals. *European Journal of Disorders of Communication*, *28*(1), 13–21.

Günzburger, D. (1995). Acoustic and perceptual implications of the transsexual voice. *Archives of Sexual Behavior*, *24*(3), 339–348.

Guzick, D. (1998). Polycystic ovary syndrome: Symptomatology, pathophysiology, and epidemiology. *American Journal of Obstetrics and Gynecology*, *179*(6), S89–S93.

Hansbury, G. (2005). The middle men: An introduction to the transmasculine identities. *Studies in Gender and Sexuality*, *6*(3), 241–264.

Heath, R. A. (2006). *The praeger handbook of transsexuality: Changing gender to match mindset.* Praeger.

Herek, G. M., Gillis, J. R., & Cogan, J. C. (2009). Internalized stigma among sexual minority adults: Insights from a social psychological perspective. *Journal of Counseling Psychology*(56), 32–43.

Hirano, M., Yoshida, Y., Yoshida, T., & Tateishi, O. (1985). Stroboscopic video recording of vocal fold vibration. *Annals of Otology, Rhinology and Laryngology*, *94*, 588–590.

Hirsch, S. (2006). Resonance. In R. K. Adler, S. Hirsch, & M. Mordaunt (Eds.), *Voice and communication therapy for the transgender/transsexual client: A comprehensive clinical guide* (pp. 207–231). Plural Publishing.

Hollien, H. (1987). "Old voices": What do we really know about them? *Journal of Voice*, *1*(1), 2–17.

Hollien, H. (1990). *The acoustics of crime.* New York: Plenum Press.

Houde, J. F., & Jordan, M. I. (1998). Sensorimotor adaptation in speech production. *Science*, *279*(5354), 1213–1216.

Houde, J. F., & Jordan, M. I. (2002). Sensorimotor adaptation of speech I: Compensation and adaptation. *Journal of Speech, Langanguage and Hearing Ressearch*, *45*(2), 295–310.

Hudson, A., & Holbrook, A. (1982). Fundamental frequency characteristics of young Black adults: Spontaneous speaking and oral reading. *Journal of Speech and Hearing Research*, *25*(1), 25–28.

Hudson, T., de Jong, G., McDougall, K., Harrison, P., & Nolan, F. (2007). F0 statistics for 100 young male speakers of Standard Southern British English. In J. Trouvain & W. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences, 6-10 August 2007, saarbrücken* (pp. 1809–1812). Saarbrücken, Germany.

Jones, J. A., & Munhall, K. G. (2002). The role of auditory feedback during phonation: studies of Mandarin tone production. *Journal of Phonetics*, *30*(3), 303–320.

Jones, J. A., & Munhall, K. G. (2003). Learning to produce speech with an altered vocal tract: the role of auditory feedback. *The Journal of the Acoustical Society of America*, *113*(1), 532–543.

Jones, J. A., & Munhall, K. G. (2005). Remapping auditory-motor representations in voice production. *Current Biology*, *15*(19), 1768–1772.

Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *The Journal of the Acoustical Society of America*, *108*(3), 1252.

Jordan-Young, R. M. (2010). *Brain storm: The flaws in the science of sex differences.* Harvard University Press.

Jurik, A. G. (1984). Ossification and calcification of the laryngeal skeleton. *Acta Radiologica: Diagnosis*, *25*(1), 17–22.

Killian, J. (1999). A description of vocal maturation among fifth- and sixth-grade boys. *Journal of Research in Music Education*, *47*(4), 357–369.

Killian, J. N., & Wayman, J. B. (2010). A descriptive study of vocal maturation among male adolescent vocalists and instrumentalists. *Journal of Research in Music Education*, *58*(1), 5 –19.

King, J. B., Lindstedt, D. E., Jensen, M., & Law, M. (1999). Transgendered voice: Considerations in case history management. *Logopedics Phoniatrics Vocology*, *24*(1), 14.

Kulick, D. (1999). Transgender and language: A review of literature and suggestions for the future. *Journal of Lesbian and Gay Studies*, *5*(4), 605–622.

Kulick, D. (2000). Gay and lesbian language. *Annual Review of Anthropology*(29), 243–85.

Kulick, D. (2003). No. *Language and Communication*, *23*(2), 139–51.

Künzel, H. J. (1987). *Sprechererkennung. Grundzüge forensischer Sprachverarbeitung.* Heidelberg, Germany: Kriminalistik Verlag.

Künzel, H. J. (1989). How well does average fundamental frequency correlate with speaker height and weight? *Phonetica*(46), 117–125.

Labov, W., Ash, S., & Boberg, C. (2005). *Atlas of North American English: Phonetics, phonology and sound change.* Mouton de Gruyter.

Ladd, D. R. (1997). *Intonational phonology.* Cambridge University Press.

Lass, N. J., Tecca, J. E., Mancuso, R. A., & Black, W. I. (1979). The effect of phonetic complexity on speaker race and sex identifications. *Journal of Phonetics*, *7*(2), 105–118.

Leden, H. von, & Moore, P. (1961). Vibratory pattern of the vocal cords in unilateral laryngeal paralysis. *Acta Oto-Laryngologica*, *73*(5), 541–550.

Lehto, L., Laaksonen, L., Vilkman, E., & Alku, P. (2008). Changes in objective acoustic measurements and subjective voice complaints in call center Customer-Service advisors during one working day. *Journal of Voice*, *22*(2), 164–177.

Levon, E. (2007). Sexuality in context: Variation and the sociolinguistic perception of identity. *Language in Society*(36), 533-554.

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, H. S. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*(54), 358–368.

Lindh, J. (2006). *Preliminary descriptive F0-statistics for young male speakers* (Vol. 52).

Linville, S. E. (1998). Acoustic correlates of perceived versus actual sexual orientation in men's speech. *Folia Phoniatrica et Logopaedica*, *50*(1), 35–48.

Liu, H., & Larson, C. R. (2007). Effects of perturbation magnitude and voice F0 level on the pitch-shift reflex. *The Journal of the Acoustical Society of America*, *122*(6), 3671–3677.

Livia, A. (2002). The future of Queer Linguistics. In K. Campbell-Kibler, R. Podesva, S. Roberts, & A. Wong (Eds.), *Language and sexuality: Contesting meaning in theory and practice* (pp. 87–99). Stanford, CA: CSLI Publications.

MacDonald, E. N., Goldberg, R., & Munhall, K. G. (2010). Compensations in response to real-time formant perturbations of different magnitudes. *The Journal of the Acoustical Society of America*, *127*(2), 1059–1068.

Maniwa, K., Jongman, A., & Wade, T. (2009). Acoustic characteristics of clear spoken English fricatives. *The Journal of the Acoustical Society of America*(125), 3962–3973.

Meier, S. C. (in-press). Sexual orientation and clinical health outcomes among female to male transsexuals.

Mercaitis, P. A., Peaper, R. E., & Schwartz, P. A. (1985). Effect of danazol on vocal pitch: A case study. *Obstetrics and Gynecology*, *65*(1), 131–135.

Mikkola, M. (2008). Feminist perspectives on sex and gender. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2008 ed.).

http://plato.stanford.edu/archives/fall2008/entries/feminism-gender/.

Miller, M. (2002). *The complete idiot's guide to music theory.* Alpha.

Mohr, J. J., & Fassinger, R. E. (2000). Measuring dimensions of lesbian and gay male experience. *Measurement and Evaluation in Counseling and Development*(33), 66-90.

Moonwomon-Baird, B. (1997). Toward the study of lesbian speech. In A. Livia & K. Hall (Eds.), *Queerly phrased: Language, gender, and sexuality* (pp. 202–214). Oxford University Press, USA.

Mount, K. H., & Salmon, S. J. (1988). Changing the vocal characteristics of a post-operative transsexual patient: A longitudinal study. *Journal of Communication Disorders*, *21*(3), 229–238.

Munhall, K. G., MacDonald, E. N., Byrne, S. K., & Johnsrude, I. (2009). Talkers alter vowel production in response to real-time formant perturbation even when instructed not to compensate. *The Journal of the Acoustical Society of America*, *125*(1), 384–390.

Munson, B. (2010). Variation, implied pathology, social meaning, and the 'gay lisp': a response to Van Borsel et al. (2009). *Journal of Communication Disorders*, *43*(1), 1–5.

Munson, B., & Babel, M. (2007). Loose lips and silver tongues, or, projecting sexual orientation through speech. *Language and Linguistics Compass*, *1*(5), 416–449.

Munson, B., McDonald, E. C., DeBoe, N. L., & White, A. R. (2006). The acoustic and perceptual bases of judgments of women and men's sexual orientation from read speech. *Journal of Phonetics*, *34*(2), 202–240.

Mupparapu, M., & Vuppalapati, A. (2005). Ossification of laryngeal cartilages on lateral cephalometric radiographs. *The Angle Orthodontist*, *75*(2), 196–201.

Murry, T., Brown, W. S., & Morris, R. J. (1995). Patterns of fundamental frequency for three types of voice samples. *Journal of Voice*, *9*(3), 282–289.

Neraud, B., & Dewailly, D. (2010). Drug-induced hyperandrogenism. In R. Azziz, J. E. Nestler, & D. Dewailly (Eds.), *Androgen excess disorders in women* (pp. 121–127). Lippincott Williams & Wilkins.

Newman, D., & Forbes, K. (1993). The effects of danazol on vocal parameters – is an objective prospective study needed? *The Medical Journal of Australia*, *19*(8), 575.

Niedzielski, N., & Koops, C. J. (2011). *The Houston Urban English Survey.* Linguistic survey database.

Nittrouer, S. (1995). Children learn separate aspects of speech production at different rates: Evidence from spectral moments. *The Journal of the Acoustical Society of America*, *97*(1), 520–530.

Nolan, F., McDougall, K., de Jong, G., & Hudson, T. (2006). A forensic phonetic study of 'dynamic' sources of variability in speech: The DyViS project. In *Proceedings of the 11th sst conference* (pp. 13–18). Auckland.

Oates, J. M., & Dacakis, G. (1997). Voice change in transsexuals. *Venereology*, *10*(3), 178–87.

Pattie, M. A., Murdoch, B. E., Theodoros, D., & Forbes, K. (1998). Voice changes in women treated for endometriosis and related conditions: the need for comprehensive vocal assessment. *Journal of Voice*, *12*(3), 366–371.

Pedersen, M. (2008). *Normal development of voice in children: Advances in evidence-based standards.* Springer.

Pegoraro-Krook, M. I., & Castro, V. C. (1994). Normative speaking fundamental frequency (SFF) characteristics of Brazilian male subjects. *Brazilian Journal of Medical and Biological Research*, *27*(7), 1659–1661.

Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *The Journal of the Acoustical Society of America*, *24*(2), 175–184.

Pierrehumbert, J. B., Bent, T., Munson, B., Bradlow, A. R., & Bailey, J. M. (2004). The influence of sexual orientation on vowel production. *The Journal of the Acoustical Society of America*, *116*(4), 1905-1908.

Podesva, R., Roberts, S., & Campbell-Kibler, K. (2002). Sharing resources and indexing meaning in the production of gay styles. In K. Campbell-Kibler, R. Podesva, S. Roberts, & A. Wong (Eds.), *Language and sexuality: Contesting meaning in theory and practice* (pp. 175–191). Stanford, CA: CSLI Publications.

Podesva, R., Roberts, S., & Campbell-Kibler, K. (2006). Sharing resources and indexing meaning in the production of gay styles. In D. Cameron & D. Kulick (Eds.), *The language and sexuality reader* (pp. 141–151). New York, NY: Routledge.

Pronovost, W. (1942). An experimental study of methods for determining natural and habitual pitch. *Speech Monographs*, *9*, 111–123.

Purcell, D. W., & Munhall, K. G. (2006). Adaptive control of vowel formant frequency: Evidence from real-time formant manipulation. *The Journal of the Acoustical Society of America*, *120*(2), 966–977.

Quain, J., Sharpey-Schäfer, E. A., Symington, J., & Bryce, T. H. (1914). *Quain's elements of anatomy.* Longmans, Green, & Co.

Queen, R. (2002). A matter of interpretation: The 'future' of 'Queer Linguistics'. In K. Campbell-Kibler, R. Podesva, S. Roberts, & A. Wong (Eds.), *Language and sexuality: Contesting meaning in theory and practice* (pp. 69–87). Stanford, CA: CSLI Publications.

Raj, S. (2011). 'Bodies in New Territories': Mapping masculinity, gender performativity and FTM embodiment in Jamison Green's 'Becoming a Visible Man'. *Altitude: An e-journal of emerging humanities work*, *9*.

Rogers, H., Smyth, R., Jacobs, G., Jensen, J. T., & van Herk, G. (2003). Vowel reduction as a cue to distinguishing gay- and straight-sounding male speech. In *Proceedings of the 2001 Annual Conference of the Canadian Linguistic Association.* University of Ottawa, Ottawa: Ottawa Papers in Linguistics.

Schilt, K. (2004). "What's to tell:" FTMs, going stealth and workplace disclosure. San Francisco, CA.

Schilt, K. (2010). *Just one of the guys? transgender men and the persistence of gender inequality.* Chicago, ILL: University of Chicago Press.

Schroeder, M., Atal, B., & Hall, J. (1979). Optimizing digital speech coders by exploiting masking properties of the human ear. *The Journal of the Acoustical*

Society of America(66), 1647–1652.

Shiller, D. M., Sato, M., Gracco, V. L., & Baum, S. R. (2009). Perceptual recalibration of speech sounds following speech motor learning. *The Journal of the Acoustical Society of America*, *125*(2), 1103–1113.

Simpson, A. P. (2001). Dynamic consequences of differences in male and female vocal tract dimensions. *The Journal of the Acoustical Society of America*, *109*(5), 2153–2164.

Simpson, A. P. (2002). Gender-specific articulatory-acoustic relations in vowel sequences. *Journal of Phonetics*, *30*(3), 417–435.

Simpson, J., & Weiner, E. (Eds.). (1989). *The Oxford English Dictionary* (2nd ed.). Oxford University Press, USA.

Smyth, R., Jacobs, G., & Rogers, H. (2003). Male voices and perceived sexual orientation: An experimental and theoretical approach. *Language in Society*, *32*(3), 329-350.

Smyth, R., & Rogers, H. (2002). Phonetics, gender, and sexual orientation. In *Proceedings of the Annual Conference of the Canadian Linguistics Association* (pp. 299–301). Montreal: University du Quebec au Montreal.

Söderpalm, E., Larsson, A., & Almquist, S. Åke. (2004). Evaluation of a consecutive group of transsexual individuals referred for vocal intervention in the west of Sweden. *Logopedics Phoniatrics Vocology*, *29*(1), 18–30.

Stemple, J. C., Glaze, L. E., & Gerdeman, B. K. (2000). *Clinical voice pathology: Theory and management* (3rd ed.). Cengage Learning.

Taylor, E. (2010). Cisgender privilege: On the privileges of performing normative gender. In K. Bornstein & S. B. Bergman (Eds.), *Gender outlaws: The next generation* (pp. 268–273). Berkeley, CA: Seal Press.

Titze, I. R. (1994). *Principles of voice production.* Englewood Cliffs, NJ: Prentice-Hall.

Tjomov, V., & Gershuni, G. (1971). A model to describe the results of psychoacoustical experiments on steady-state stimuli. In (p. 36-49). Leningrad: Nauka.

Traunmüller, H. (1990). Analytical expressions for the tonotopic sensory scale. *The Journal of the Acoustical Society of America*(88), 97–100.

Traunmüller, H., & Eriksson, A. (1995). *The frequency range of the voice fundamental in the speech of male and female adults.* `http://www2.ling.su.se/staff/hartmut/`.

van Bezooijen, R. (1995). Sociocultural aspects of pitch differences between Japanese and Dutch women. *Language and Speech*, *38*(3), 253–265.

van Bezooijen, R., & Gooskens, C. (1999). Identification of language varieties. *Journal of Language and Social Psychology*, *18*(1), 31–48.

van Borsel, J., De Bruyn, E., Lefebvre, E., Sokoloff, A., De Ley, S., & Baudonck, N. (2009). The prevalence of lisping in gay men. *Journal of Communication Disorders*, *42*(2), 100–106.

van Borsel, J., De Cuypere, G., Rubens, R., & Destaerke, B. (2000). Voice problems in female-to-male transsexuals. *International Journal of Language and Communication Disorders*, *35*, 427–442.

van Borsel, J., De Cuypere, G., & van den Berghe, H. (2001). Physical appearance and voice in male-to-female transsexuals. *Journal of Voice*, *15*(4), 570–575.

van Mersbergen, M. R., Verdolini, K., & Titze, I. R. (1999). Time-of-day effects on voice range profile performance in young, vocally untrained adult females. *Journal of Voice*, *13*(4), 518–528.

Villacorta, V. M., Perkell, J. S., & Guenther, F. H. (2007). Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception. *The Journal of the Acoustical Society of America*, *122*(4), 2306–2319.

Waksler, S. (2001). Pitch range and women's sexual orientation. *Word*, *52*, 69–77.

Walls, N. E., & Costello, K. (2010). "Head ladies center for teacup chain": Exploring cisgender privilege in a (predominantly) gay male context. In S. Anderson & V. Middleton (Eds.), *Explorations in diversity: Examining privilege and oppression in a multicultural society* (2nd ed., pp. 81–93). Belmont, CA: Brooks / Cole.

Wardle, P. G., Whitehead, M. I., & Mills, R. P. (1983). Non-reversible and wide ranging voice changes after treatment with danazol. *British Medical Journal*, *287*(1), 946–946.

Weglarski, A., Sewall, A., Schiavetti, N., Metz, D. E., & Whitehead, R. L. (2000). Effect of vowel environment on consonant duration: An extension of normative data to adult contextual speech. *Journal of Communication Disorders*, *33*(1), 1–10.

Wells, J. C. (1982). *Accents of English: Volume 1.* Cambridge, UK: Cambridge

University Press.

Wong, A., Roberts, S., & Campbell-Kibler, K. (2002). The future of Queer Linguistics. In K. Campbell-Kibler, R. Podesva, S. Roberts, & A. Wong (Eds.), *Language and sexuality: Contesting meaning in theory and practice* (pp. 1–21). Stanford, CA: CSLI Publications.

Zimman, L. (2009). "The other kind of coming out": Transgender people and the coming out narrative genre. *Gender and Language*, *3*(1), 53–80.

Zimman, L. (2010). Female-to-male transsexuals and gay-sounding voices: A pilot study. In (Vol. 22, pp. 1–21). Boulder: University of Colorado.

Zraick, R. I., Birdwell, K. Y., & Smith-Olinde, L. (2005). The effect of speaking sample duration on determination of habitual pitch. *Journal of Voice*, *19*(2), 197–201.

Zraick, R. I., Gentry, M. A., Smith-Olinde, L., & Gregg, B. A. (2006). The effect of speaking context on elicitation of habitual pitch. *Journal of Voice*, *20*(4), 545–554.

Zraick, R. I., Nelson, J. L., Montague, J. C., & Monoson, P. K. (2000). The effect of task on determination of maximum phonational frequency range. *Journal of Voice*, *14*(2), 154–160.

Zraick, R. I., Skaggs, S. D., & Montague, J. C. (2000). The effect of task on determination of habitual pitch. *Journal of Voice*, *14*(4), 484–489.

Zwicker, E., & Terhardt, E. (1980). Analytical expressions for critical band rate and critical band-width as a function of frequency. *The Journal of the Acoustical*

Society of America(68), 1523–1525.